

Copyright
by
Emily Karen Weisburst
2018

The Dissertation Committee for Emily Karen Weisburst
certifies that this is the approved version of the following dissertation:

**Essays on the Economics of Law Enforcement
Institutions and Policy**

Committee:

Sandra Black Youngblood, Supervisor

Jason Abrevaya

Leigh Linden

Becky Pettit

**Essays on the Economics of Law Enforcement
Institutions and Policy**

by

Emily Karen Weisburst

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2018

Essays on the Economics of Law Enforcement Institutions and Policy

Publication No. _____

Emily Karen Weisburst, Ph.D.
The University of Texas at Austin, 2018

Supervisor: Sandra Black Youngblood

This dissertation consists of three chapters on the economics of law enforcement institutions and policy. In the first chapter, I examine the importance of individual police officers to arrest outcomes in interactions with civilians. I show that the likelihood of an arrest is not only a function of incident timing, geography, offense type, and other contextual factors but also critically depends on the identity of the police officer who responds to a call for service. Examining detailed data on more than 1,850 police officers responding to over 160,000 calls for service from the Dallas Police Department, I find that officers vary widely in their arrest behavior, with a 1 standard deviation increase in an officer's propensity to arrest resulting in a 33% increase in the likelihood that a given incident results in an arrest.

In the second chapter, I investigate the impact of police hiring on crime rates in municipalities in the U.S. In this chapter, I use a novel esti-

mation approach to which exploits variation in federal Community Oriented Policing Services (COPS) hiring grants, while also controlling for the endogenous decisions of police departments to apply for these grants. Using data from nearly 7,000 municipalities, I find that a 10% increase in police employment rates reduces violent crime rates by 13% and property crime rates by 7%. The results also provide suggestive evidence that law enforcement leaders are forward-looking.

In the third chapter, I explore the impact of police on student discipline and academic outcomes. This chapter provides the first causal estimate of funding for school police on student outcomes, leveraging variation in federal Community Oriented Policing Services (COPS) grants. Exploiting detailed data on over 2.5 million students in Texas, I find that funding for police in public schools results in a small but significant reduction in high school graduation and college enrollment.

Table of Contents

Abstract	iv
List of Tables	x
List of Figures	xiii
Chapter 1. “Whose Help is on the Way?”	
The Importance of Individual Police Officers in Law Enforcement Outcomes	1
1.1 Introduction	2
1.2 Institutional Background and Description of Data	8
1.2.1 Dallas Police Department Data	8
1.2.2 Protocols for Dispatch and Call Response	10
1.3 Empirical Model	14
1.4 Summary Statistics	19
1.5 Importance of Individual Officers and Dispersion in Officer Behavior	21
1.5.1 Baseline Results	21
1.5.1.1 Tests of Officer Sorting to Incidents	25
1.5.2 Officer Demographics	30
1.5.3 Robustness Tests	32
1.5.3.1 Bootstrap Tests of Primary Results	32

1.5.3.2	Robustness Specification Tests	33
1.6	Racial Bias Among Officers	35
1.6.1	Testing for Racial Bias Using Information on Officers and Arrestee Outcomes	35
1.6.2	Results of the Test of Racial Bias	39
1.6.2.1	Baseline Racial Bias Results	39
1.6.2.2	Robustness of Racial Bias Results	42
1.7	Conclusion	46
1.8	Tables and Figures	49
Chapter 2.	Safety in Police Numbers: Evidence of Police Effectiveness from COPS Grant Applications	68
2.1	Introduction	68
2.2	COPS Hiring Grant Program	74
2.3	Empirical Model and Data Sources	79
2.4	Results	86
2.4.1	Summary Statistics	86
2.4.2	Primary Results	87
2.4.3	Model Validity and Treatment Effects Over Time	92
2.4.4	Robustness Tests	95
2.4.5	Mechanisms: Crime Clearances	100
2.4.6	Shifts in Policing Focus	103
2.5	Conclusion	105
2.6	Tables and Figures	109

Chapter 3. Patrolling Public Schools:	
 The Impact of Funding for School Police on Student Discipline and Long-term Education Outcomes	121
3.1 Introduction	122
3.2 Background and Institutional Context	127
3.2.1 Federal COPS Grants for School Police	127
3.2.2 School Resource Officers and School Discipline in Texas	130
3.3 Empirical Model and Data Sources	133
3.4 Results	141
3.4.1 Summary Statistics	141
3.4.2 Baseline Results	144
3.4.3 Treatment Heterogeneity	148
3.4.3.1 Heterogeneous Effects across Race and Income .	148
3.4.4 Robustness Tests of the Baseline Model	153
3.5 Conclusion	157
3.6 Tables and Figures	159
Appendices	171
Appendix A. Appendix:	
 “Whose Help is on the Way?”	
 The Importance of Individual Police Officers in Law Enforcement Outcomes	172
A.1 Appendix Tables and Figures	173
A.2 Coefficients in the First Stage of Model	189
A.3 Empirical Bayes Shrinkage Estimates	193

A.4 Economic Model for Racial Bias Test	199
A.5 Data Appendix	207
Appendix B. Appendix:	
Safety in Police Numbers:	
Evidence of Police Effectiveness from COPS Grant Applications	213
B.1 Appendix Tables and Figures	214
B.2 Data Appendix	219
Appendix C. Appendix:	
Patrolling Public Schools:	
The Impact of Funding for School Police on Student Discipline and Long-term Education Outcomes	223
C.1 Appendix Tables and Figures	224
C.2 Data Appendix	234
Bibliography	239
Vita	254

List of Tables

1.1	Summary Statistics	49
1.2	Summary Results and Tests of Sorting, Comparison to Unlikely Response Sample	57
1.3	Officer Effects and Officer Demographics	58
1.4	Robustness Specification Tests	60
1.5	Racial Bias Test, Officer and Arrestee Race	62
2.1	Summary Statistics, Covariates and Outcomes	111
2.2	Grant Characteristics	112
2.3	Impact of Police on Crime	113
2.4	Impact of Police on Crime, Specific Crimes	114
2.5	Crime Clearances as a Mechanism, Reported vs. Cleared Specific Crimes	117
2.6	Robustness Tests and Samples, Specific Crimes (Part 1)	118
2.7	Shifts in Policing Focus, Arrests for Specific Crimes	120
3.1	Summary Statistics	160
3.2	Summary Statistics, Grant Data Weighted by Student Observations	161
3.3	Short-term Student Discipline Outcomes	162
3.4	Expanded Short-term Student Outcomes	163
3.5	Long-term Student Outcomes	164

3.6	Short-term Student Discipline Outcomes, by Demographic Group	165
3.7	Expanded Short-term Student Outcomes, by Demographic Group	166
3.8	Long-term Student Outcomes, by Demographic Group	167
3.9	Robustness Tests, Short-term Student Discipline Outcomes	169
3.10	Robustness Tests, Long-term student outcomes	170
A.1	Summary Statistics, Officer Sorting Robustness Samples	175
A.2	Additional Policing Outcomes	177
A.3	Officer Effects and Officer Demographics, Additional Policing Outcomes	179
A.4	Officer Effects and Officer Demographics, Robustness Specifications	180
A.5	Summary Statistics, Racial Bias Test Sample	181
A.6	Racial Bias Test, Officer Characteristics in First Stage	183
A.7	Replication of Racial Bias Tests in Prior Literature	188
A.8	Covariate Coefficients in First Stage of Arrest Model	192
B.1	Impact of Police on Crime, Influence of Model Type and Covariates	214
B.2	Impact of Police on Index I Arrests, By Demographic Group	217
B.3	Impact of Police on Other Arrests, By Demographic Group	218
C.1	Grant Effects on School District Budget Data	226
C.2	Effects by COPS Grant Type	227
C.3	Effects by Grantee Type, School District Police Department vs. Other Grantees	228

C.4	Table 3.6 with Main Effects: Short-term Student Discipline Outcomes, by Demographic Group	231
C.5	Table 3.7 with Main Effects: Expanded Short-term Student Discipline Outcomes, by Demographic Group	232
C.6	Table 3.8 with Main Effects: Long-term Student Outcomes, by Demographic Group	233

List of Figures

1.1	Relative Proportion of Total R^2 , Components of Model	51
1.2	Distribution of Officer Effects, $\hat{\theta}_i$	52
1.3	Tests of the Importance of Officer Sorting to Officer Effect Dis- tribution	53
1.4	Bootstrap Benchmark Test Distribution	59
1.5	Distribution of Officer Effects across Officer Race	64
2.1	COPS Grant Program over Time	109
2.2	Timing of Treatment, Application and Acceptances	115
2.3	Placebo Timing Tests, Randomized Attribution of Grant Ap- plications and Acceptances (1,000 Replications)	116
3.1	COPS Grant Funding for School Police in Texas	159
3.2	Timing of Grant Acceptance and Application Treatment Effects on Student Discipline	168
A.1	Steps involved in an Incident Response	173
A.2	Police Beats and Police Divisions in Dallas, TX	174
A.3	Power of Racial Bias Test, Bootstrap Simulation	185
A.4	Adjusted and Unadjusted Officer Effects	198
B.1	Timing of Treatment, Specific Crimes, Application and Accep- tances	215

C.1	Depiction of Model Identification, Comparison of Two Hypothetical Districts	224
C.2	Discipline Grant Treatment Effects, by Grade	225
C.3	Middle School: Timing of Grant Acceptance and Application Treatment Effects on Student Discipline,	229
C.4	High School: Timing of Grant Acceptance and Application Treatment Effects on Student Discipline,	230

Chapter 1

“Whose Help is on the Way?”

The Importance of Individual Police Officers in Law Enforcement Outcomes

The public’s perception of police fairness is essential to the willingness of citizens to cooperate with the police and is fundamental to establishing police legitimacy. However, little is known about whether police officers are actually fair and impartial in their application of the law. In this paper, I show that the likelihood of an arrest is not only a function of incident timing, geography, offense type, and other contextual factors but also critically depends on the identity of the police officer who responds to a call for service. Examining detailed data on more than 1,850 police officers responding to over 160,000 calls for service from the Dallas Police Department, I quantify variation in arrest behavior across individual police officer decision-makers and relate this variation to officer demographic and employment characteristics. I find that police officers are important determinants of arrest outcomes, with the variation in individual officer behavior accounting for 10-15% of the explainable variation in arrests. Officers vary widely in their arrest behavior, with a 1 standard deviation increase in an officer’s propensity to arrest resulting in a

33% increase in the likelihood that a given incident results in an arrest. Additionally, I find limited evidence that officer differences are driven by racial bias, a result that may be related to an array of progressive police reforms that have been adopted by the Dallas Police Department in recent years.

1.1 Introduction

The public’s perception of police fairness is essential to the willingness of citizens to cooperate with the police and is fundamental to establishing police legitimacy [98, 39].¹ In recent years, public trust in police has become increasingly strained; recent survey evidence finds that less than 30% of individuals in high-crime and low-income areas believe that police “make fair and impartial decisions in the cases they deal with” or that police “make decisions based on the law and not their personal opinions or beliefs” [67]. Distrust in police is related to the perception that officers are racially biased; according to national polling, 69% of Black and 54% of Hispanic Americans think that “police tend to target minorities,” compared to only 29% of White Americans [93]. In 2015, American confidence in police officers reached its lowest point in more than 20 years, driven by high profile police use-of-force incidents and shootings [61]. In the ensuing debate, pundits have often made conflicting claims about

¹Survey based research in criminology finds that when civilians believe that they have been treated fairly by the police, a concept often termed “procedural justice,” they may be more likely to cooperate with police officers, and this cooperation may have benefits for public safety [98, 39].

the causes of these high profile police incidents, sometimes asserting that “any officer would have responded in the same way” and at other times claiming that the events are “isolated incidents attributable to ‘bad actors’ that do not reflect the rest of a department.”²

However, little is known about whether police officers are actually fair and impartial in their application of the law. On a more basic level, there is scant evidence of whether officer decisions actually matter to the outcomes of police interactions after considering the context of an incident. Further, if police officer decisions do matter, little is known about how much officers differ from each other in their actions. If there are differences in police officer behavior, how large are these differences? Is officer arrest behavior characterized by racial bias, and if so, how important is racial bias in explaining officer differences?

Managers in police departments face a similar principal-agent problem as managers in firms; they are impacted by the behavior of individual officers but cannot fully control officers’ actions given limited resources. Understanding the trade-offs between alternative monitoring policies is an important area of study, particularly in the setting of policing, where there are

²This phrasing is not directly attributed to any single pundit or public figure. An example of the first argument can be found in opinion pieces through the organization *Blue Lives Matter*, which was established as a reaction to the *Black Lives Matter* movement [8]. The second argument was recently invoked by Attorney General Sessions as a reason to cease the Department of Justice’s enforcement of consent decree agreements with police departments, which were established to address civil rights concerns related to law enforcement actions [26].

large potential consequences for both public safety and civilians that interact with the police. This paper measures the scope of individual police officers to impact law enforcement outcomes, a necessary first step to clarifying these trade-offs.

Police officers often operate in the field, alone or in small groups, and have substantial legal latitude in their conduct and response to different situations. At the same time, police departments are increasingly incorporating technology and data to standardize operations, potentially limiting the importance of individual officer decisions in police work.³ The ability of police officers to invest effort differently across incident types can be a productive form of police discretion when police resources are limited and there is a trade-off between exerting effort on serious and non-serious crimes. However, within particular incident types, behavioral differences across officers are more likely to result from differences in officer skills, experience, and preferences. This study is the first to estimate the degree and importance of police discretion across officers, conditional on incident context.

I analyze police officer behavior using a sample of over 160,000

³Criminologists have long noted that police work is characterized by discretion, with researchers adopting a broad definition of discretion that encompasses variation in work-related decisions, interpretation and implementation of the law, and the use of extra-legal factors, such as suspect race, in decision-making [85, 39, 75]. Technological advancements have increased police reliance on data for surveillance of suspects, tracking and monitoring of police employee activities, automation of reporting, and focusing patrol on areas with high offending rates, or crime “hot spots.” Advancements in technology have the potential to exacerbate differences in police treatment of civilians, or could serve to reduce police discretion [10, 58].

calls for service (or 9-1-1 calls) and approximately 1,850 police officers from the Dallas Police Department in Texas. I estimate the contribution of individual officers to predicting arrests, controlling for detailed information on the characteristics of incidents, including call urgency and dispatch code, complainant characteristics, and time and geographic factors. I then isolate each police officers' propensity to make arrests and estimate the degree of dispersion in this propensity across officers. Throughout the analysis, I pay particular attention to patterns of officer sorting to particular incident types, and conduct a number of robustness checks to verify that the observed dispersion in arrest behavior across officers is not due to selection.

I find that a 1 standard deviation increase in an officer's permanent arrest propensity results in a 33% increase in the likelihood that any given incident results in an arrest, suggesting there is substantial variability in police officer responses to similar incidents. Further, this variation in individual police officer behavior accounts for approximately 10-15% of the explainable variation in incident arrest outcomes. Observable officer characteristics, including salary, experience, age, race and gender, account for only a small portion of the variation in officer arrest behavior.

In economics, research on police decision-making has largely focused on measuring racial bias in police traffic stops [e.g. 56, 4, 5, 46, 6, 48, 65]. New work by [40] extends this literature to officer decisions to use violent force. Collectively, this literature has found mixed evidence that police officers ex-

hibit taste-based preferences for racial discrimination, with results that vary by study setting and the test used to detect racial bias. While the literature has frequently exploited aggregate officer demographic characteristics, nearly all of the work in this space does not incorporate officer identity in measuring racial bias.⁴

I find limited evidence of taste-based racial bias among officers, despite the fact that I document large differences in total arrest behavior across officers. I adapt a test of racial bias proposed in [6] to the regression framework in my study and fail to find systematic patterns of racial bias. Moreover, officer race is not a particularly important factor in explaining differences in officer behavior. These results may be attributable to progressive police reforms that have been adopted by the Dallas Police Department in recent years, including implicit racial bias training, de-escalation training, and the use of body-worn cameras.

In addition to providing the first estimate of officer-level police discretion, this paper makes several other contributions. First, the analysis in this study uses responses to calls for service that are originated by civilian complainants, a setting that researchers have not exploited to study police decision-making.⁵ I am able to explicitly measure the importance of officers

⁴An exception is the working paper by [47], that measures individual officer specific effects in its test of racial bias applied to officer decisions to issue speeding tickets.

⁵To the best of my knowledge, this is the first study that uses high frequency call data to study police officer behavior and decisions. Prior papers using 9-1-1 call data have examined a number of other topics, including the the impact of high profile use-of-force incidents on

versus incident context because the incident setting is given at the time of the response and is therefore not manipulable by responding officers. With the exception of [104], who studies racial bias of state troopers who are randomly dispatched to motor vehicle accidents, researchers in economics have typically restricted their attention to interactions that are initiated by police officers, such as general traffic stops, stop and frisk interviews, and speeding tickets. Importantly, these interactions are a choice variable of the police officers involved. A growing body of research finds that race can also be a factor in police decisions to make traffic (or pedestrian) stops, and that studies that focus only on the outcomes of traffic stops neglect to consider police discretion that contributes to sample selection [56, 46, 48]. A major advantage of using call for service data to study police discretion is that each observed incident is originally initiated by a complainant and not by a police officer.

An additional contribution of this study is the number and variety of policing outcomes that I examine. Unlike traffic stops, call for service data provides a diverse cross-section of police work, allowing examination of responses to different types of crime and incidents. The data allows measurement of outcomes related to suspect identification, arrest charge severity, and time spent by officers responding to incidents. Detailed demographic informa-

complainant crime reporting [27], forecasts of temporal and spatial patterns of 9-1-1 call activity [17], police patrol optimization in hot-spots [100], and factors that affect response time to calls for service [69]. [12] use a similar data set from the Dallas Police Department covering calls between 2000-2007 to study whether increases in crime cause future crime, at the neighborhood level. [102] also uses Dallas Police Department data from 2009 on 911 calls and patrol car location to examine the impact of police presence on crime rates.

tion on officers, arrestees, and civilian complainants also provide rich controls in the model and enable tests of racial bias.

The paper proceeds as follows. Section 1.2 summarizes institutional features of the Dallas Police Department and the data sets used in this study. Section 1.3 describes the empirical model and estimation approach. Section 3.4.1 provides summary statistics of the data and setting. Section 1.5 details the findings related to the importance of individual officers in determining arrest outcomes and the dispersion in officer arrest behavior. Section 1.6 extends the analysis to test for the presence of racial bias among officers. Section 1.7 concludes.

1.2 Institutional Background and Description of Data

1.2.1 Dallas Police Department Data

The setting for this study is the Dallas Police Department in Texas. Dallas is a large and diverse urban center, with over 1.2 million residents and a population that is 42% Hispanic, 24% Black and 29% White [14].⁶ Crime rates in the city of Dallas are similar to other cities of its size in the U.S., with 694 violent crimes and 3,440 property crimes per 100,000 residents in 2015 [36]. Given its large population, the Dallas Police Department responds to between 4,000 - 7,000 calls for service each month.

⁶I capitalize Black, Hispanic, and White for stylistic consistency throughout the paper.

In recent years, the Dallas Police Department (DPD) has become a model of police reform following changes spear-headed by Chief David O. Brown (2010-2016). DPD’s reform efforts have included increasing officer training requirements in de-escalation techniques and racial bias in policing, employing body cameras, and firing some of its most poorly performing officers [52, 97]. Though Dallas shares several characteristics with other large cities across the country, the progressive nature of the department is relevant when interpreting the results of this study.

To improve accountability and transparency in the department, DPD joined the Obama Administration’s Police Data Initiative in 2015 and released a number of data sets on its operations. This project uses public data released by DPD covering responses to reported incidents, records of persons involved in incidents (suspects and complainants), and arrests between December 2014 and mid-October 2017.⁷

Each call incident can be linked to a location, dispatch code, the time a call was placed, and the time and duration of the call response. Incidents are then linked to complainants, suspects and arrestees, using data that includes names and demographics of each of these groups. Most important for this project, the data also includes the names and badge numbers of officers

⁷Given that data transparency is a relatively new DPD initiative, there is limited documentation describing the data releases. To interpret the data sets, I worked with personnel at DPD to better understand the data sets used in this project as well as the department protocols relevant to police responses to calls for service.

who respond to calls for service.

The DPD incident data is a partially restricted subset of the universe of call for service incidents. For privacy reasons, the DPD data files exclude records for sexually-oriented offenses, offenses involving juveniles, and offenses involving social service referrals.⁸ I use a liberal definition of arrests, coding an incident as having an arrest if any of the files indicate that an arrest was made. I supplement the DPD data sets with demographic information on police officers obtained through an Open Records Request to the city of Dallas. The Open Records Request data includes officer race, gender, salary, and job title (see Data Appendix [C.2](#) for more detail).

1.2.2 Protocols for Dispatch and Call Response

When a civilian calls DPD for police assistance, they are connected to a 9-1-1 call-taker. The call-taker then creates an active call report that summarizes important facts related to the incident, including the incident location, and relevant descriptions of the events. Active call reports also include a dispatch code that categorizes the incident/offense type.

⁸Further, the raw incident files represent records of all incidents reported to the police, and include some incidents that do not originate with calls from civilian complainants. To address this issue, I use information on the identity of the complainant and the dispatch code to exclude records that are likely to originate with a police officer rather than a complainant, or are unlikely to be initiated through a phone call to DPD. Examples of these exclusions include calls for service to assist an officer on another call or incident and calls originated by the Dallas Police Department or City of Dallas (listed as complainant).

Given a set of open active calls, DPD dispatchers then work with police officers to assign available officers to incidents. Calls are dispatched according to their priority, or the level of severity and urgency of the incident. When there is a long call queue, responses to low priority calls are postponed until more serious calls have been resolved. The pool of available officers when a call is received depends on patrol responses to other incidents at the time. Figure A.1 depicts the steps involved in responding to a call for service in Dallas.

Patrol officers are the primary responders to calls for service. Officers are assigned to work in 1 of 7 police divisions in the city for 8-hour shifts, or watches, from 12am-8am, 8am-4pm, or 4pm-12am.⁹ Regular patrol shift schedules are set once a year, based on the seniority of officers.¹⁰ Depending on the needs of the department, officers may also choose to work overtime patrol shifts outside of their regular shift schedules, though these shifts are also set in advance, typically a month or a week prior.¹¹ With minimal exceptions, calls

⁹Patrol officers are assigned to work in particular police sectors, a sub-area of a division, under a sector sergeant (there are 35 sectors in total). Depending on the sector, patrol officers may also be assigned to work in a beat within the sector during a shift (234 beats in total). Despite sector assignments, patrol officers may be called to respond to any call within their division. The three 8-hour shifts listed are approximate, in practice some officers work 10 hour shifts and other officers have start and end times that are slightly staggered across police “watches.”

¹⁰Officers can bid to change their shift assignments each December. Changes are assigned based on officer seniority and take effect at the end of January or beginning of February in the following year. New officers that are still in field training or have recently graduated from the police academy, “rookies,” are not eligible to petition for shift changes.

¹¹“Call-answering” overtime initiatives are designed to address anticipated increases in call volume, most often in the summer months. Sometimes these patrol initiatives may

are assigned to patrol officers who work within the geographic police division where the call incident occurred.

Officers typically conduct patrol in police cars, alone or in pairs. At the beginning of each shift, officers may choose to patrol with another officer, depending on the number of cars available for that shift. Each car is considered an “element” that can be dispatched to an incident. Throughout a shift, paired officers respond to all calls together.

If more than one patrol element is available to respond to an incident at the time of dispatch, dispatchers consider a number of factors in their assignment of available officers. More serious incidents may require or benefit from a response by multiple officers. Additionally, officers who are geographically close to an incident are more likely to be dispatched to the incident, especially if the call is urgent. At the same time, depending on availability, officers may volunteer to take particular calls as they are posted. If officers choose to respond to calls based on incident characteristics that are unobservable and high (low) arrest officers choose incidents with a high (low) likelihood of arrest, then officer sorting will lead to an upward bias in the estimate of dispersion in officer arrest behavior. To address this concern, I conduct a series of tests to verify that officer sorting does not affect the empirical estimates (see Section [1.5.1.1](#)).

work to target a particular offense, such as burglary. In some cases, overtime shifts may be announced only a couple of days in advance, but the typical lead time is a week or a month.

When the assigned patrol element arrives at the scene of the incident, the responding officer(s) gather information, investigate the scene and assist the complainant or victim. In serious incidents, assisting officers may be called to the scene, through the request of the responding officer or dispatcher, or because an additional officer volunteers to participate in the call.¹² First responders are responsible for coordinating other officers that may arrive at the scene as well as filing the incident report when the response has concluded.

After the first responder files the incident report, it is submitted to a staff reviewer at DPD who examines the incident report for completeness. After this initial review, the incident may be assigned to a detective in an investigative unit based on the offense type. The assigned detective will then pursue additional investigation of the incident if warranted.

Over the course of an incident response, officers determine whether a criminal offense has occurred and may identify a suspect and/or make an arrest. Alternatively, an arrest may be made at a later date after a detective takes over responsibility for a follow-up investigation of the incident. A suspect is identified in 17% of incidents, and 9% of incidents result in an arrest. Individual responding officers have the ability to influence arrests directly, by making the decision to apprehend an individual at the scene of the incident,

¹²If the responding officer is patrolling alone, this officer is designated as the first responder, and if there are two responding officers, one officer in the pair is designated as the first responder. Typically, one officer in a pair is designated as the first responder for the duration of the shift.

or indirectly, by laying the groundwork for an investigation by gathering information for the incident report. In practice, most arrests do not involve a prolonged follow-up investigation and the responding officer is typically involved in the arrest.¹³

1.3 Empirical Model

I focus on two metrics to assess the importance of individual officers in arrest outcomes. First, I measure the collective importance of individual officers' arrest propensities in predicting arrest outcomes. Second, I estimate each officer's permanent arrest propensity and measure the dispersion in permanent arrest propensity across officers. The first measure summarizes how important individual officer decisions are relative to the context of an incident, while the second measure captures how different officers are from one another.

As a first step, I estimate the following linear probability model,

$$Arrest_{ikgt} = \theta_i + \theta_j + \pi X_{kt} + \delta_{gt} + \phi_g + \varepsilon_{ikgt}$$

where i indexes the responding officer, j indexes a co-responding officer (if present), k indexes the incident, g indexes geographic location, and t indexes

¹³When there is information on the time of arrest and arresting officer, 93% of arrests occur within a day of the response to the call and 95% of arrests involve the original responding officer. The first rate is imputed from direct information on arrest officers and arrest dates when available, as well as the data upload date for an arrest record when not available. Information on the officers that executed an arrest is available for approximately 40% of arrests in the data.

time. The outcome $Arrest_{ikgt}$ is the primary focus of the analysis and denotes whether an arrest was made in association with the incident. X_{kt} are a set of incident specific characteristics, including 15 aggregated dispatch codes or incident type categories and 11 location type codes (e.g. street or residence). X_{kt} also contains complainant characteristics, including the number of complainants, whether there was a victim injury, and the race and gender of complainants.¹⁴ Further, X_{kt} includes indicators for the number of hours that have passed since a patrol shift began. Importantly, the model also controls for the *urgency or severity* of the call, defined as the number of minutes that pass between when a call occurs and the time of dispatch. This urgency measure enters the model as both a linear and quadratic term, to account for non-linear relationships between call urgency and arrests.

I include controls for the police beat where an incident occurred, ϕ_g , to control for time-constant differences in arrest patterns that are location specific. There are 234 beats in Dallas and each is fully contained within 1 of the 7 police divisions in the city.

I also include a set of shift specific controls, δ_{gt} , to capture time-varying location-specific arrest patterns that are associated with specific shift assignments. These variables are Police Division* Day-of-the-Week* 8-hour Shift* Month* Year fixed effects. To increase power, I do not include a separate

¹⁴For calls with multiple complainants, I define each complainant variable using the maximum value for the complainant group, allowing complainants to have multiple races and genders.

indicator for each individual shift, but rather aggregate them into month by year groups. For example, the four Tuesday evening shifts in the Central Division are grouped in January 2016. These shift variables flexibly control for unobservables at the shift level that vary over time.¹⁵

θ_i measures the time-invariant or permanent arrest propensity of officer i . Given the numerous controls in the empirical model, θ_i represents an officer specific effect that is measured within incident type, shifts, and geographic location. In cases when there are two responding officers, I include a control for the identity of the other responding officer, θ_j . Observations with two responders are duplicated, so that each officer gets a record of participating in the incident response through the θ_i term. This procedure allows measurement of individual officer effects net of the effects of a co-responder, using a model similar to prior work on peer effects in production [e.g. 94, 74]. In this way, the specification addresses omitted variable bias related to police officer decisions to pair with another officer, as well as potential direct effects

¹⁵Because of the large number of fixed effects in the model, I employ a fixed effect estimation algorithm. In the base model, there are 217,633 observations, 1,851 first officer categories, 2,337 second officer categories, 5,143 shift categories and 234 beats (after excluding singletons). I use an algorithm developed in [23] to estimate the set of fixed effects indirectly through an iterative procedure that provides a point estimate value for each fixed effect. Rather than estimating the model's fixed effects by including corresponding indicator variables as controls in the model, this algorithm effectively initializes each fixed effect within a fixed effect group, and then iterates the estimation until both the sum of squared residuals is minimized and the coefficient on each set of fixed effect terms is 1. This procedure is programmed in the STATA command `reghdfe`, and is notable among similar algorithms for its fast computation time. This algorithm shares features with other procedures that have been used to estimate multiple high-dimensional fixed effects [e.g. 44, 45, 50]. I estimate all sets of fixed effects in the model jointly in this way, or θ_i , θ_j , δ_{gt} and ϕ_g .

attributable to the co-responder.¹⁶ I restrict the sample to observations where the officers responding have at least 25 incident records to improve precision in the estimation of θ_i .¹⁷

Using this model, I calculate the two metrics of importance in this study. First, I measure the importance of the θ_i terms as predictors in the model, by calculating the proportion of explainable variation that is attributable to these parameters. I do this by estimating the R^2 (and Adjusted R^2) from the full model and the model without θ_i and θ_j terms included. I then calculate the proportion of the total R^2 (and Adjusted R^2) given by the including officer fixed effects: $(R^2_{total} - R^2_{w/oFE})/R^2_{total}$. I interpret this ratio as the relative importance of individual officers in explaining the model variation in arrest outcomes.

Second, I calculate the dispersion in officer-level permanent arrest propensity as the standard deviation of the distribution of θ_i across officers. In

¹⁶The results in this paper are robust to two alternate formulations that exclude fixed effect terms for co-responders, θ_j . These alternate specifications are (1) observations that are duplicated when there are co-responders but exclude co-responder fixed effects and (2) observations that are not duplicated when there are co-responders but only consider effects from the first listed responder. Formulation (2) further restricts the sample and excludes usable observations for officers, limiting the number of officer effects estimates from 1,851 to 1,692. The preferred model officer effects are highly correlated to the officer effects from alternate formulations with correlations of 0.77-0.85. The baseline specification is preferred to these alternate formulations because it allows both responders to contribute to the outcome and more accurately measures individual officer effects by adjusting for potential contributions of co-responders.

¹⁷This restriction excludes 4% of incidents in the sample, but allows estimation of fixed effects to be based on a reasonable number of observations per officer. 5% of co-responder incidents include only records for only one responder given this restriction. Further limiting the sample to exclude these “one-sided” incidents does not affect the results.

order to establish a conservative estimate of police officer dispersion, I adjust the estimates of θ_i terms using Empirical Bayes techniques.¹⁸ Broadly, Empirical Bayes methods consider each observation of officer arrest propensity as a noisy estimate of each officer’s true arrest propensity and shrink those estimates that are less precise towards the expected value of the distribution officer arrest propensity across all officers. In this setting, the expected value of the distribution of officer arrest propensity is zero because officer effects are defined in relative terms (given that a constant is included in the model). This procedure both improves precision in the estimates of the θ_i terms and establishes a conservative estimate of dispersion in police officer arrest behavior.

I calculate the adjusted estimates of θ_i using the following steps. First, I consider the model above as a first stage and construct a composite residual term $\hat{r}_{ikgt} = \hat{\theta}_i + \hat{\varepsilon}_{ikgt}$, and an average officer residual, $\bar{r}_i = \frac{1}{N_i} \sum_{N_i} \hat{r}_{ikgt}$. This residual is estimated using a model that includes officer fixed effects in the first stage to allow for arbitrary correlations between responding officers and the other covariates in the model, in a manner similar to [16]. Using these residuals, I calculate the adjusted officer arrest propensity using the following transformation:

$$\begin{aligned}\hat{\theta}_i^{EB} &= \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma_{\varepsilon,i}^2}{N_i}} \cdot \bar{r}_i \\ \sigma_A^2 &= \sigma_r^2 - \sigma_\varepsilon^2\end{aligned}$$

¹⁸Empirical Bayes shrinkage estimates are often used in the economics of education literature measuring teacher value added [e.g. 66, 16, 62, 1].

where σ_r^2 is computed using the sample analog of the average squared composite residual and σ_ε^2 is the average squared within officer composite residual, each calculated from the first stage model. The “shrinkage factor,” $\sigma_A^2/(\sigma_A^2 + \frac{\sigma_{\varepsilon,i}^2}{N_i})$, adjusts officer arrest propensity toward zero when the number of observations per officer, N_i , is small, or the variation in the officer effect, $\sigma_{\varepsilon,i}^2$, is large. The $\hat{\theta}_i^{EB}$ values represent a “posterior” distribution of officer effects, correcting for sampling noise in the estimation (see Appendix A.3 for more detail). I calculate dispersion in officer arrest behavior as the standard deviation of these adjusted officer effects. Dispersion in the adjusted officer effects will be smaller than dispersion in the unadjusted effects because the adjustment shrinks noisier estimates towards the mean of the distribution. Throughout this paper, I focus on results using these adjusted estimates, and refer to these adjusted estimates as $\hat{\theta}_i$. The results are not an artifact of this adjustment though and are qualitatively similar when unadjusted fixed effects from the first stage are used.

1.4 Summary Statistics

Tables 1.1.A and 1.1.B summarize the data used in the analysis. The first column covers the total sample at the incident-level, while the second column covers the analysis sample, which restricts the sample to officers with 25 or more incident responses and duplicates observations with two responding officers. The analysis sample includes over 215,000 observations for 1,851

officers.

Black, White, and Hispanic complainants are each involved in a third of the incidents in the sample, with an average number of 1.6 complainants per incident. Black and White complainants are overrepresented relative to the Dallas population, which is over 40% Hispanic. In contrast, White patrol officers respond to 50% of incidents, while Black and Hispanic officers respond to 25% and 20% of incidents, respectively. Approximately a third of arrest incidents involve arrestees who are Black, 10% who are Hispanic, and 10% who are White, when proportions are calculated relative to total arrests. Relative to the sample of incidents with demographic information for arrestees, 48% of arrestees are Black, 16% are White, and 16% are Hispanic.¹⁹ White officers and Black arrestees are over-represented in police incidents relative to the population of Dallas.

The average officer arrest rate as a portion of his responses is 11%. Approximately 5% of incident responses involve a police officer in training, 2% involve a police sergeant, and 15% involve a female officer. Over 76% of incident arrests include an arrest for a misdemeanor offense and a third of incident arrests include a felony arrest. Averaged across incident responses, DPD patrol officers earn approximately \$60,000 per year. About a third of incident responses involve two responding officers.

¹⁹Demographic information is not available for each arrest in the data, and covers approximately 70% of arrests.

On average, it takes 24-25 minutes for a patrol officer to be dispatched to an incident after a call is made, with a standard deviation of 28 minutes. The variation in this dispatch time lag highlights the fact that dispatchers prioritize calls based on their severity and that officers cannot immediately respond to all incidents. The most common dispatch codes are for burglaries of motor vehicles and residences. At the time of dispatch, only a small number of incidents are designated as known violent offenses; robberies, criminal assaults, armed encounters, and active shootings collectively comprise less than 10% of incidents. A victim is injured in less than 10% of incidents.

Overall, the samples summarized in Tables [1.1.A](#) and [1.1.B](#) are very similar. The only material difference is mechanical; the analysis sample has a larger number of observations with 2 responding officers, because these incidents are duplicated in this sample. This consistency suggests it is suitable to generalize results in the analysis sample.

1.5 Importance of Individual Officers and Dispersion in Officer Behavior

1.5.1 Baseline Results

I find that the context of a call for service is relatively more important in predicting an arrest than the identity of the officer that responds to the call. However, individual officers are also a significant determinant of arrest outcomes. As discussed earlier, I measure the collective importance of

individual officers to predicting arrest outcomes of incidents, by comparing the proportion of variation explained with and without officer fixed effects in the model of arrest outcomes.

Figure 1.1 shows the contribution of different controls to the total model R^2 and Adjusted R^2 . I estimate these relative proportions sequentially from the bottom bar to the top, first adding the patrol shift fixed effects and indicators for hours passed since the beginning of a shift (δ_{gt} and hour indicators in X_{kt}), then the police beat location fixed effects (ϕ_g), followed by the dispatch call types and call severity variables, location type variables, complainant characteristics (each components of X_{kt}), and lastly the officer fixed effects (θ_i and θ_j). Each percentage is calculated as an additional contribution of R^2 to the total or: $(R^2_{currentbar} - R^2_{priorbar})/R^2_{total}$. The explanatory power of the officer fixed effects is therefore the relative contribution of these controls after controlling for all other variables in the model.

I find that the officer fixed effects account for 11.9% of the explainable variation measured using Adjusted R^2 and 16.5% of explainable variation measured using R^2 . Factors specific to an incident, including dispatch code, call severity, location type and complainant characteristics, account for 57 - 69% of the explainable variation in the model. Geography and time variables associated with an incident account for 19 - 26% of the model variation. A detailed discussion of the first stage covariate estimates, X_{kt} , can be found in Appendix A.2.

Because the order of adding variables matters to calculating R^2 , I add individual officer effects to the model last to ensure that the estimate of officer importance is as conservative as possible. The total R^2 and Adjusted R^2 of the model is 0.23 and 0.2, respectively, suggesting that officer effects explain 2.4 - 3.8% of the total variation in arrests. When officer fixed effects are the only controls in the model, they can explain 6.2 - 8% of the total variation in arrests.

Next, I show that individual police officers vary substantially in their arrest behavior. Figure 1.2 shows the estimated distribution of officer effects, $\hat{\theta}_i$, calculated using the procedure described in Section 1.3. For each officer, $\hat{\theta}_i$ represents his/her permanent or time-invariant arrest propensity, conditional on time and geography controls and incident, location, and complainant characteristics. This estimated distribution has a longer right tail, showing that a small number of officers have especially high arrest propensities.

Swapping an officer that has a low arrest propensity with one that has a high arrest propensity can critically change the outcome of a call response. Given that 10.9% of sample observations result in an arrest, with a standard deviation of 0.311, a 1 standard deviation in $\hat{\theta}_i$ corresponds to 0.11 standard deviations in the total arrest outcome. Additionally, a 1 standard deviation increase in officer arrest propensity corresponds to 0.23 standard deviations in the variation in arrest outcomes predicted by the model. In per-

centage terms, a 1 standard deviation increase in an officer’s arrest propensity results in a 32.8% increase in the likelihood that a given incident results in an arrest. Further, moving from the 10th to 90th percentile in the officer fixed effect distribution translates to a 76% increase in arrest probability.

Given of the richness of the data used in this project, I am able to extend the analysis to a number of other policing outcomes of interest. Table A.2 extends the analysis to seven additional outcomes: identification of a suspect, arrest for a felony offense, arrest for a misdemeanor offense, arrests made by an officer that responds to an incident alone, arrests in urgent and non-urgent calls (defined by splitting the sample at the median call urgency), and the amount of time spent by an officer responding to an incident. Officers responding to calls are involved in identifying the appropriate suspect and making an arrest of the culpable party. Individual officers are important to the determination of each of these outcomes, and there is substantial variation across individual officers in the likelihood that each outcome occurs.

The officer effects, $\hat{\theta}_{i,y}$, estimated for each of the additional outcomes, y , are positively correlated with the $\hat{\theta}_i$ distribution for the base arrest outcome, with correlation coefficients of 0.55 - 0.84.²⁰ An exception is the officer effect for time spent responding to an incident, which is virtually uncorrelated with arrest propensity, $\hat{\theta}_i$. Panel (B) of Table A.2 shows that officer

²⁰Additionally, the officer effects estimated for a given officer across all observations, θ_i , and as a co-responder, θ_j , are strongly related with a correlation of 0.71.

effects for felony and misdemeanor arrests are also uncorrelated, implying that these actions are trade-offs and that officers with high misdemeanor arrest propensities will not necessarily exhibit high felony arrest propensities. Likewise, officer effects across urgent and non-urgent call settings are not strongly related. Aside from these relationships, officer effects derived from these different outcomes are largely positively correlated. Extending the analysis to these additional policing outcomes shows that officers exhibit heterogeneity across different dimensions of police work.

1.5.1.1 Tests of Officer Sorting to Incidents

As discussed above, the fact that patrol officers can choose to respond to certain incidents creates a potential concern. While I control for an array of observable characteristics, if officers systematically sort to calls based on the unobservable characteristics of incidents, the estimates of individual officer arrest propensity could be biased. Two sorting patterns appear plausible, first, officers who have a high arrest propensity may volunteer for incidents with a high likelihood of arrest, and second, officers who have a low arrest propensity may volunteer for incidents with a low likelihood of arrest. In either case, the estimates of $\hat{\theta}_i$ will be positively correlated with the error terms ε_{ikgt} and inflate the dispersion in officer fixed effects, $\hat{\theta}_i$.

To address this concern, I conduct four tests of officer sorting in the analysis. First, I focus two different settings where officer sorting is less likely

to impact the results. If sorting is driving the results, estimates of dispersion should be smaller in settings where sorting is less likely. The first setting I consider consists of incidents that are dispatched when few officers are available to respond to calls, so officers have less choice in their responses. I define this “Low Availability” sub-sample by counting the number of unavailable officers that are responding to other incidents at the time of each incident, using information on the time officers spend responding to prior incidents. I then keep incidents that have more unavailable officers than the median value within a patrol shift*month cell, δ_{gt} , to account for variation in total staffing across shifts.

In the second setting, I focus on incidents where an officer’s observed response to a call is *unlikely* to have occurred, and are therefore less susceptible to concerns about officer sorting. I define this “Unlikely Response” sub-sample using techniques similar to propensity score matching. First, I estimate a linear probability model for each officer i that predicts whether i responded to a particular incident, conditional on the full set of covariates in the model, including incident characteristics, X_{kt} , geographic beats, ϕ_b , shift indicators, δ_{gt} , and fixed effects for co-responders, θ_j . From these regressions, I obtain a predicted response probability for each officer across all observations. I then code observations as *unlikely* responses if the predicted response probability is below its median value among the actual incident responses for each officer. By construction, this sub-sample of *unlikely* responses consists of incident responses that are not characterized by predictable sorting of officers.

I restrict each robustness sample to officers with at least 25 observations in the relevant sub-sample, yielding 1,504 officers in the “Unlikely Response” sample and 1,556 officers in the “Low Availability” sample. I then compare estimates within these sub-samples to corresponding samples that include all observations for these officers. Tables [A.1.A](#) and [A.1.B](#) show that the characteristics of the robustness samples are similar to the primary analysis sample.

Figure [1.3\(a\)](#) and [1.3\(b\)](#) show the results of restricting the observations to the “Low Availability” and “Unlikely Response” sample. The graphs show a close match in the distributions across each of the samples and their corresponding baselines. The correlation in officer effects is 0.785 between each of the robustness samples and their corresponding baseline.

Table [1.2](#) shows that the estimated dispersion in officer effects is comparable across the analysis sample, the “Low Availability” sample, and the “Unlikely Response” sample. First, the proportion of variation explained by the fixed effects is similar across the main sample and the “Low Availability” and “Unlikely Response” samples, accounting for 11 - 17% of the explainable variation in the baseline (columns (1), (2) and (4)) and 12 - 19% of the explainable variation in the robustness samples (columns (3) and (5)). Second, a 1 standard deviation increase in officer fixed effects increases the probability of arrest by 32.8% in the primary sample, 38.5% in the “Low Availability” sample, and 41.5% in the “Unlikely Response” sample. If dispersion in offi-

cer behavior is increased by officer sorting, we would expect the estimates of dispersion to be lower in these robustness samples than the baseline. However, we observe the opposite pattern here. The larger dispersion estimates in the sorting robustness samples is in part due to lower precision of the officer effects estimates in these smaller samples. Across each of these settings, the officer effects are highly correlated and the estimates of officer dispersion are qualitatively similar.

Next, I consider how important incident characteristics are in the estimation of $\hat{\theta}_i$. I calculate the correlation between the distribution of $\hat{\theta}_i$ from the full model to $\hat{\theta}'_i$ estimated from a model that omits incident characteristics that could influence an officer's decision to respond to a call, X_{kt} and ϕ_g . If officer arrest propensity is orthogonal to observable incident characteristics, these distributions will be perfectly correlated. This test is informative if unobservable incident characteristics are correlated with observables, an assumption that is often applied in tests of endogeneity. Figure 1.3(c) shows that the $\hat{\theta}'_i$ distribution is somewhat more disperse than the base $\hat{\theta}_i$ estimates that include incident characteristics; however, the estimates across these distributions have a correlation of 0.906. This high correlation suggests that incident characteristics are not very important controls in the estimation of the officer effects distribution.²¹

²¹I have also measured the joint significance of incident and geographic characteristics that officers may use to choose responses, X_{kt} and ϕ_g , in predicting officer effects, θ_i . I do this by duplicating θ_i to the incident-level, regressing these effects on the full model specification, and taking a joint F-test of X_{kt} and ϕ_g . The F-statistic of this test is 1.79, and is significant

I also construct a balance test that asks whether high and low arrest officers are similarly likely to respond to incidents, given their characteristics. To do this, I create a composite index of covariates in the model as the predicted arrest likelihood of an incident, estimated from a model that excludes officer effects. Next, I separately estimate the likelihood that an officer in the top 25% (high arrest officers) and bottom 25% (low arrest officers) responds to an incident, by regressing an indicator variable for whether officers in these groups responded to an incident on covariates in the model, again excluding officer effects. In Figure 1.3(d), I plot the difference in the response likelihood for high and low arrest officers in relation to the predicted arrest likelihood of incidents. I find a nearly flat relationship between these variables with a fitted slope of 0.1, implying that high and low arrest officers are approximately equally likely to respond to high or low arrest incidents.

Overall, these tests suggest that the baseline estimates of officer dispersion are not biased by officer sorting. This pattern is likely related to the rich set of controls in the model, which include geography, time, patrol shift controls, and complainant characteristics, as well as a proxy for the urgency or severity of each incident.

but not particularly large (with standard errors clustered at the focal officer, i , and shift level, δ_{gt} , to account for duplicated incident records and officer effect outcomes). Together with the test above that finds a very high correlation between officer effects estimated with and without observables, this test suggests that incident characteristics have a small but significant relationship to θ_i .

1.5.2 Officer Demographics

How does officer arrest propensity relate to officer demographics? A natural next step is to consider how the estimated officer fixed effects, $\hat{\theta}_i$, are associated with officer demographic characteristics.²² I regress $\hat{\theta}_i$ terms on officer race, gender, age, trainee or sergeant status, experience, and experience squared in Table 1.3.²³ The officer demographics regressions offer information about whether officers with specific traits systematically differ in their arrest propensities.

Officers with more experience have higher arrest propensities, while officer trainees, and female officers have lower arrest propensities. All else equal, the likelihood of arrest is 7% higher when a responding officer has 10 years of experience instead of 5 years of experience, 9% lower when a responding officer is a police trainee, and 4% lower when an officer is female. While Hispanic officers are not statistically different from the omitted other race officer group, Hispanic officers have statistically lower arrest propensities than White or Black officers. Relative to White officers, officers that are Hispanic

²²So far, I have not included officer demographics in the estimation model because these factors do not vary over time and are collinear with the officer fixed effect terms. While some officer characteristics like salary, trainee status, or age, may in fact vary over time, they change slowly relative to the timing of incident observations. Given their low frequency, these changes would be absorbed by other time factors in the model if included directly in the estimation equation. Further, the officer demographic data was obtained as a snapshot file and does not include records of promotions or salary changes.

²³Salary is omitted from this regression because it is nearly perfectly correlated with experience, given the compensation formulas used by the department. Results are similar when salary is used instead of years of experience.

have 6% lower likelihood of arrest on average. Aside from this coefficient for Hispanic officers, race is not a systematically important factor in predicting total officer arrest propensity. The regressions are very similar using officer effects derived from the “Low Availability” and “Unlikely Response” samples.

There are a number of interesting patterns present when this analysis is extended to officer effects from the set of additional policing outcomes in Tables A.3. Again, police trainees are less likely to make felony and misdemeanor arrests and arrests in non-urgent calls. Officers with higher levels of experience have higher propensities to identify suspects and make arrests when responding to calls alone. Older officers spend more time responding to incidents than younger officers. Hispanic and Black officers are less likely to have high propensities to identify suspects and are more likely to spend a longer time responding to calls for service. As in the general arrest outcome, female officers have lower felony arrest propensities and Hispanic officers are least likely to make arrests across each of these additional settings.

Overall, however, demographic factors do not explain a large share of the total variation in officer effects, with regression R^2 statistics of 0.02 - 0.04. Instead, this analysis shows that the substantial variation in arrest behavior observed across officers is due to unobservable characteristics of police officers, such as officer preferences or unobservable dimensions of productivity.

1.5.3 Robustness Tests

1.5.3.1 Bootstrap Tests of Primary Results

How large is the dispersion in officer arrest behavior? One way to assert that the distribution of officer arrest propensities is meaningful is to benchmark the observed standard deviation in officer effects, $\hat{\theta}_i$, to the amount of variation that would be observed across officers if there were no “true” officer effects. Even in the absence of officer differences, there will be some measured variation in outcomes across officers, simply due to idiosyncratic variation in the error term or “noise.”

To confirm that the results in this study reflect actual variation in behavior across officers, I next construct a wild cluster bootstrap test to benchmark the results [13]. In this test, I calculate the residuals, \hat{r} , and predicted outcome values, \hat{Arrest} , from a regression that does not include officer fixed effects, thereby assuming a null hypothesis that the true value of all θ_i and θ_j terms is zero and outcomes are not systematically correlated within officer observations. While imposing this null hypothesis, the procedure allows errors to be correlated within shift clusters, δ_{gt} , to account for common shocks within geographic areas, time periods, and officer groups, while also accounting for error correlation across duplicated incident observations. In each bootstrap iteration, I apply weights of $w \in \{-1, 1\}$ to residuals \hat{r} that are constant within the observations for each shift cluster δ_{gt} and assigned with equal probability for each shift cluster. Using these weights, I reconstruct a

new outcome variable as the predicted outcome plus the weighted residual, $\tilde{Arrest}_b = \hat{Arrest} + w_b \hat{r}$. I then regress this new outcome variable, \tilde{Arrest}_b , on the fully specified model that includes officer fixed effects and calculate the dispersion of the Empirical Bayes adjusted officer effects for each iteration. I also recover the R^2 and Adjusted R^2 contribution of officer fixed effects for each iteration of the bootstrap. The distribution of these metrics across bootstrap iterations provides a benchmark that corresponds to variation in officer effects that would arise if there were no “true” officer effects simply because of idiosyncratic variation or “noise.”

Figure 1.4 displays the results of these tests using 250 bootstrap replications. These graphs confirm that the estimated variation in officer effects and the contribution of the officer effects to explaining the model variation is not due to noise in the data. Each of the model estimates are well outside the 95% confidence interval given by the bootstrap tests.

1.5.3.2 Robustness Specification Tests

Next, I consider several alternate specifications of the base model in Table 1.4. In column (2), I substitute the 234 police beat categories with narrower geographic area controls of 1,143 police reporting areas in the city of Dallas. Column (3) alternatively substitutes the police beat controls with geographic controls that vary by time, or 1,225 Police Sector*Month category variables. Column (4) includes 21,358 individual 8-hour shift level indicators

(Date*8-hour shift*Division) instead of the monthly aggregated shift indicators, δ_{gt} , used in the primary specification. By conditioning on individual 8-hour shifts within police divisions, this specification absorbs variation in arrests at the date by geography level, accounting for factors such as changing weather conditions, holidays, and other day specific events in the city of Dallas. In column (8), I replace the 15 aggregated dispatch codes and 11 aggregated location codes used in the main specification with 117 specific dispatch code and 34 specific location type categories in the raw incident data.²⁴ Across specifications these additional controls do very little to change the analysis, offering estimates that are close to the base model. In fact, the correlation between the base distribution of officer fixed effects and these alternate specifications is above 0.95.

Columns (6) and (7) consider alternative procedures to adjust the estimates for precision instead of the Empirical Bayes method used in the primary results. In column (6), I report the dispersion in unadjusted officer fixed effects from the first stage, where the sample is restricted to officers with more than 100 observations. In column (7), I weight the unadjusted officer fixed effects by the number of observations per officer and calculate a weighted standard deviation as the dispersion metric. Across both of these alternative precision methods, the dispersion estimates are similar to the base model, with a 1 standard deviation in officer effects corresponding to a 38%

²⁴I collapse dispatch code categories to increase the estimation speed of the model and address small cell categories in the majority of the analysis.

increase in arrest probability. Lastly, in column (8), I report dispersion in officer behavior using the unadjusted first stage officer fixed effects. With no adjustment for precision, this standard deviation estimate is moderately larger than the base model, corresponding to a 46% increase in arrest probability.

The baseline results are robust to a number of different alternate specifications and validity tests. In sum, individual officers vary substantially in their arrest behavior, and individual officers are important in determining the outcomes of police interactions.

1.6 Racial Bias Among Officers

1.6.1 Testing for Racial Bias Using Information on Officers and Arrestee Outcomes

Is there evidence of racial bias among police officers in Dallas? In this section, I test for the presence of racial bias by investigating relationships between officer race and arrestee race.

I adapt the test of racial bias used in work by [6] examining officer bias in traffic stops. [6] develop a model of officer interactions that allows officers of different races to have differing arrest behavior as long as these differences are independent of suspect race. Specifically, they test whether the relative rank order of arrest rates across officer race groups is the same within each suspect race. For example, if White officers are more likely than Black officers to arrest a suspect from any race group, then this reflects the

total arrest preferences of White and Black officers but does not imply that either group is racially biased. Alternatively, if Black officers have higher arrest rates than White officers for White suspects *and* White officers have higher arrest rates than Black officers for Black suspects, either Black officers or White officers are racially biased (or both). Critically, the test does not find evidence of racial bias if officers *statistically discriminate* against suspects, or if officers of all races use suspect race as a signal of offending characteristics that are correlated with race. More generally, the test allows suspect race groups to have different compositions or unobservable characteristics that may cause differences in arrest rates for suspects of different races *across all officers*. Arrest rates can differ across suspect race groups in the test because differences in officer race arrest rates are always measured as a relative ranking within a suspect race group.

Appendix [A.4](#) presents the economic framework for the racial bias test, adapted from [6] to the call for service setting. In the call for service setting, officers maximize their expected benefit of making an arrest and face costs of exerting effort that may differ by officer and suspect race. Officers make effort choices after viewing the suspect race and a signal of whether the arrest is *feasible* or whether there will be a sufficient basis for an arrest if the officer exerts effort. The distribution of signals of feasibility and underlying feasibility of arrests are allowed to differ by suspect race, and these differences enable officers to statistically discriminate against suspects of a particular race. A key difference between the call for service setting and the traffic stop setting

is that I do not directly observe suspect race for all incidents.²⁵ Instead, I execute this test by comparing officer arrest propensities for different arrestee race outcomes.

I consider unconditional arrestee race outcomes that take a value of 1 if an individual is arrested and is a particular race, either Black, Hispanic, or White, and 0 otherwise. As before, I estimate $\theta_{i,r}$ terms for each arrestee race outcome, r , using the model outlined in Section 1.3. Next, I regress the $\theta_{i,r}$ terms on the full set of officer demographic characteristics and examine the impact of officer race on officer arrest propensity. This regression framework allows me to directly test whether the rank of arrest rates across officer race changes for different arrestee race outcomes.²⁶

The test used in this paper offers three new advantages. First, I am able to test for racial bias among officers in a setting that is not affected by officers electing to initiate interactions. Prior work studying racial bias in policing has examined interactions between officer and suspect race in officer-

²⁵I have limited information on suspect identification in the data. There are two sets of records of suspects in the data: (1) I observe suspects identified and demographic information for these suspects in a subset of cases (limited to data prior to 2017), (2) I observe demographic characteristics of suspects that are unknown to officers at the conclusion of the response. I treat both of these records as outcomes in the data because they are entered by responding officers and may be a function of officer effort. Arrest records can be considered a subset of (1) and are available for the duration of the sample period.

²⁶My preferred specification uses the two stage regression in order to be consistent with the analysis in Section 1.5.2. I also conduct the test by inserting officer demographic variables directly into the first stage while omitting the officer fixed effects in Section 1.6.2.2. The results are similar in both formulations.

initiated incidents, such as traffic stops [e.g. 56, 4, 5, 6, 48].²⁷ These papers consider suspect race as a given characteristic of a traffic stop; however, in reality, suspect race is also a choice variable of the officer, who chooses to stop a particular individual. In this paper, I am able to test for the presence of racial bias in a setting that is not affected by this form of selection because calls are initiated by complainants and not officers.

Second, I use a regression framework to control for a large array of observable incident characteristics. [6] address the fact that different officers may face different types of incidents by re-sampling their data to create comparable incident sets across officer race. In my setting, I use regression models to measure officer race arrest effects adjusted for observable differences in the composition of incidents across officers.

Lastly, I leverage officer identifiers to better understand relationships between officer behavior and arrestee race. I am able to use officer identifiers to trace the distribution of officer effects by race for each of the arrestee race outcomes and find that most variation in officer behavior is within rather than across officer race.

The racial bias test used in this paper cannot detect cases when all officers exhibit similar racial bias toward a particular group, a limitation of prior tests as well. As discussed above, Black arrestees are markedly overrepre-

²⁷An exception in this literature is [104], which studies racial bias of state troopers who are randomly dispatched to motor vehicle accidents rather than officer-initiated interactions.

sented in the sample, making up 40% of total incident arrests and only 24% of the Dallas population.²⁸ If arrests with Black suspects result in higher arrest rates of Black individuals *across all officers*, this pattern could be consistent with statistical discrimination, or institutional discrimination. Institutional racial bias will occur when the organizational priorities of the department direct resources toward policing one race group relative to others, and all officers behave similarly given these priorities. Likewise, the higher representation of Black arrestees could also be consistent with uniform police officer attitudes of taste-based racial bias against Black suspects. Recent evidence on the importance of implicit racial bias in decision-making could be consistent with uniform taste-based discrimination against minority groups [e.g. 32].

1.6.2 Results of the Test of Racial Bias

1.6.2.1 Baseline Racial Bias Results

Before turning to the results of the test of racial bias, I assess whether incident characteristics are comparable across officer races. Tables [A.5.A](#) and [A.5.B](#) show that the composition of incidents is similar for different officer race groups. The most meaningful difference is that each officer race is more likely to respond to a complainant that has the same race. Because social interactions are largely segregated, individuals are most likely to be victimized by an individual of the same race, so the variation in complainant race also

²⁸These proportions are calculated relative to total arrests, not relative to arrests with arrestee race information (70% of observations).

means that officers are more likely to respond to an incident with a suspect of their own race. The empirical strategy used to conduct the test of racial bias controls directly for complainant race to address this difference in exposure. I limit the sample used in this test to observations where responding officers have a single race and each officer has more than 25 observations.²⁹

Table 1.5 shows the results of the test of racial bias among officers. Each regression estimates the relationship between officer demographic control variables and the measured officer effect, $\theta_{i,r}$, calculated from a predictive model of whether the incident resulted in an arrest of an individual of race, r (corresponds to columns). The F-tests show whether the officer race coefficients are statistically different from one another.

In the full sample (Panel A), White officers have a higher propensity than Hispanic and Black officers to make any arrest. The rank order of White officers having the highest arrest propensity, followed by Black officers and then Hispanic officers is consistent for each arrestee race outcome, controlling for incident characteristics. Each of the links in this order are not always statistically different from one another. In column (2), White officers are statistically more likely than Hispanic officers to arrest Black suspects, but the difference between Hispanic officers and Black officers is not significant. Likewise, Black officers are more likely than Hispanic officers to arrest Hispanic

²⁹The racial bias analysis sample is constant for all arrestee race outcomes. This restriction mechanically increases the number of White officers with co-responders; responder pairs that are a single race are more likely to be White because there are more White officers overall.

suspects, but the difference between Black and White officers is not significant (column (3)). Lastly, in column (4), White officers are more likely than Black officers to arrest White suspects, however the difference between Black officers and Hispanic officers is not significant. None of the officer race coefficients is statistically different from the omitted race category of “other race” officers for any of the outcomes.

I am also able to leverage individual officer identifiers to better understand patterns of officer behavior across arrestee race. In Figure 1.5, I compare the distribution of measured officer effects, $\theta_{i,r}$, by officer race for each arrestee race outcome. The distributions of officer fixed effects by officer race are quite close to one another for each outcome. The distribution of White officer arrest effects is noticeably to the right of the distributions of Black and Hispanic officers for Black arrestee and White arrestee outcomes. Importantly, these figures do not suggest that the rank order of officer arrest effects by officer race changes across these outcomes. Moreover, the pictures suggest that there is more variation in arrest propensity *within* officer race than between officer races. In fact, if officer effects were constant across officer race, the dispersion in officer effects for each of these outcomes would decrease by less than 1%.³⁰ In other words, race is not the most important determinant

³⁰This change is calculated using coefficients from the regressions in Table 1.5. For the racial bias sample, a 1 standard deviation increase in the officer effect distribution corresponds to a 39% increase in total arrest likelihood, and a 58%, 110%, and 92% increase in Black, Hispanic, and White arrestee race outcomes, respectively. The larger percentage changes for the demographic arrestee outcomes is related to their smaller mean values.

of an officer’s arrest behavior.

The test cannot reject the null hypothesis that officers are not racially biased. This pattern could be consistent with the progressive reforms adopted by the Dallas Police Department. As discussed above, DPD enacted a number of reform initiatives prior to and during the sample period, including implicit racial bias training, de-escalation training, the use of body-worn cameras, and sharing data on its operations with the public.

1.6.2.2 Robustness of Racial Bias Results

In Panels (B) and (C) of Table 1.5, I replicate the test of racial bias using the “Low Availability” and “Unlikely Response” robustness samples to address concerns about officer sorting. In both of these alternative samples, the test cannot reject the null hypothesis that White officers have the highest arrest propensity, followed by Black officers, and then Hispanic officers. Extending the analysis to these samples suggests that the racial bias result is not driven by officer sorting. In Table A.6, I show that this finding is also robust to considering a simplified version of the test, which does not include officer fixed effects in the first stage of the model but instead inserts officer characteristics as direct controls in the first stage.

The racial bias test used in this paper has the advantage of being unlikely to yield a false positive claim of bias, an attractive feature given that a finding of racial bias is politically sensitive. However, a potential concern

about the test used in this paper is that it has low power to detect racial bias. The first reason that the test may have low power is conceptual; the test cannot detect patterns of racial bias that are shared across all officers or patterns of racial bias that are consistent with certain officer race groups arresting all types of suspects more frequently than other officer race groups. In addition to this limitation, the test may have lower power because it is a joint hypothesis test.

I empirically assess the power of the racial bias test using a bootstrap simulation. The simulation estimates the rate that the null hypothesis will be rejected when an alternative hypothesis of racial bias is imposed in the data. I simplify this analysis by imposing artificial changes to one officer race coefficient and allowing other officer race coefficients to remain unchanged. I set the altered coefficient to become larger for different-race arrestee outcomes and smaller for the same-race arrestee outcome. For example, I allow the White coefficient to increase in the second-stage regression of Black and Hispanic arrest outcomes, $\theta_{i,Black}$ and $\theta_{i,Hispanic}$, and decrease in the second-stage regression of the White arrest outcome, $\theta_{i,White}$. Specifically, I set the deviations in the White coefficient to be a constant percent change in the relevant arrestee race outcome relative to the Black and Hispanic officer average. This change imposes the alternative hypothesis that the officer race arrest rate ranking differs across arrestee race outcomes.³¹

³¹This simulation is conducted using the following steps (for the White coefficient example). Allow each officer race coefficient in the second stage to be denoted by α_r . I first

Figure A.3 plots the power of the test for different values of the alternative hypothesis, or different percent deviations in arrests caused by the altered coefficient. This illustrative exercise shows that I am able to accept an alternative hypothesis of a 14.5% deviation in arrest outcomes due to changes in the White officer coefficient with 95% power. For Black and Hispanic officer coefficients, I am able to accept corresponding alternative hypotheses of an 10.5% and 18.5% deviation with 95% power, respectively. This power exercise shows that the test is unlikely to detect particularly small differences in racial bias but that it will succeed in rejecting the null hypothesis when there are substantial patterns of bias.

Lastly, I benchmark the findings in this paper by replicating a

regress the second stage $\theta_{i,r}$ outcomes on officer demographics excluding the White officer race coefficient, $White_{i,r}$ and recover a predicted $\hat{\theta}'_{i,r}$ estimate and a predicted residual $\hat{r}_{i,r}$. In each bootstrap iteration, I draw a wild bootstrap weight $w_b \in \{-1, 1\}$ with equal probability for each weight. I then impose an alternative hypothesis on the $White_{i,r}$ coefficient to construct a simulated value for each officer effect, $\tilde{\theta}_{i,r}^b = \hat{\theta}'_{i,r} + \Delta White_{i,r} + w_b \hat{r}_{i,r}$. I set the magnitude of Δ to be a constant percent increase for Black and Hispanic arrestee outcomes and an equivalent percent decrease in the White arrestee outcome, so that the alternative hypothesis is a true difference in White officer ranking across the arrestee race outcomes. These percent changes are set relative to Black or Hispanic officer averages for the total arrest outcome, using the higher of the two officer groups (Black or Hispanic) for percent increases and the lower of the two for the percent decrease. In other words, a Δ of a 5 percent deviation will equal the $\Delta = \alpha_{Black} + 0.05 / (\text{mean}(Arrest_{Hispanic}) + \alpha_{Black})$ for the $\theta_{i,Hispanic}$ regression if Black officers were the reference category with larger coefficient, and $\Delta = \alpha_{Hispanic} - 0.05 / (\text{mean}(Arrest_{White}) + \alpha_{Hispanic})$ for the $\theta_{i,White}$ regression if Hispanic officers were the reference category with the smaller coefficient. Using these simulated bootstrap values, $\tilde{\theta}_{i,r}^b$, I regress officer effects on the full set of officer demographic variables and use F-tests to determine whether the ranking of officer race groups changes across arrestee race outcomes. Figure A.3 plots the rejection rate or power of the test for different values of the alternative hypothesis, Δ .

number of tests used in the prior literature in Table A.7.³² First, I conduct a simple comparison of the demographic composition of arrestees and residents in Dallas in Panel (A) and find that Black arrestees disproportionately represented relative to their share in the city population. Panel (B) shows that conditional arrest rates for Black and Hispanic arrestees are approximately 49%, while these rates are significantly higher for Whites at 55%. Using the framework in [65], this pattern provides evidence of racial bias of officers because average “hit rates” are higher for White suspects, suggesting that officers may “over-suspect” minority suspects and “under-suspect” White suspects. In Panel (C), I replicate an aggregate version of the test in [6] without adjusting for incident characteristics and fail to find evidence of racial bias.³³ Lastly, in Panel (D), I construct a version of the test presented in [5] which proposes that officer behavior is characterized by racial bias when officers arrest individuals of their own race at lower rates than individuals of other races. Using this test, I find that officers actually arrest individuals of their own race at *higher* rates than individuals of other races. In general, I fail to reject the null hypothesis that officers are not biased replicating tests of racial bias that

³²For simplicity, I each of the tests uses aggregated arrest statistics for different demographic subgroups in the data, that are not adjusted for incident characteristics, geographic or time factors, or officer effects. In the source papers, the authors have also conducted various regression-based tests that adjust for other factors related to a police search or incident. In the majority of the paper, I treat suspect identification as an outcome of a police response, but the comparisons in this section condition on suspect information directly. Suspect identification records are available prior to 2017.

³³Interestingly, the rank order in the raw aggregate data is different than in the regression adjusted test described above; Black officers are most likely to make arrests, followed by Hispanic officers and then White officers. This suggests that Black and Hispanic officers respond to incidents that are more likely to result in arrest.

are used in the prior literature. The data in this setting fails simpler tests of racial bias that involve stricter assumptions (Panels (A) and (B)), while more nuanced tests that apply more general assumptions do not show evidence of racial bias among officers (Panels (C) and (D)).

Overall, I do not find clear patterns of racial bias among officers. In this section, I adapted the test proposed in [6] to the call for service setting and found that the relative ranking of officer arrest propensity across officer race does not significantly differ across arrestee race outcomes. Further, the data suggests that officer arrest propensity differs more within race than across race, implying that an officer’s race is not the most important determinant of his arrest behavior.

1.7 Conclusion

Individual police officers are critical to the outcomes of police work. While the context of an incident, such as geographic location, time, and offense type, largely determines if the response to the incident will result in an arrest, individual police officer decisions also matter. Analyzing high frequency call for service data from Dallas, Texas, I find that the individual officers that respond to calls for service account for 10-15% of the explainable variation in arrest outcomes.

Police work is characterized by discretion and police officers differ

from one another. I find that a 1 standard deviation increase in officer arrest propensity increases the likelihood of an arrest by 33%. In general, observable officer demographic characteristics do not explain a large share of the variation in arrest behavior across officers.

Further, the variation in individual officer arrest propensity does not appear to be driven by racial bias. I find that the rank order of arrest propensity by officer race does not significantly change across arrestee race outcomes. Given this evidence, I am unable to detect substantive patterns of racial bias among officers. While these findings counter traditional concerns about discrimination against minority individuals in policing, the results may be related to a number of progressive police reforms that have been adopted by the Dallas Police Department, including implicit racial bias training, de-escalation training, and the use of body-worn cameras. However, more research is needed to understand how these reform initiatives may alter officer biases and actions.

Having established that individual officers are critical to the outcomes of police and civilian interactions, questions remain for future research. First, what are the marginal welfare costs (or benefits) of arrests that result from police discretion? An arrest may be a positive or negative welfare outcome depending on the incident context, culpability of the arrestee, and severity of the offense. Some arrests may serve to increase public safety through deterrence or incapacitation of the suspect, while others may have limited

public safety benefits and cause undue burden to the suspect and his/or her family. Given that the outcomes of police interactions can have varying consequences, it is unclear whether the average arrest propensity of officers is optimal, or if the ideal arrest propensity might correspond to a different point in the distribution. A greater understanding of the quality of marginal arrests could improve response protocols of police departments.

Lastly, what are the best policy levers to reduce dispersion in arrest behavior? From a fairness perspective, investing in reducing dispersion in officer behavior could yield benefits in the form of increased trust in law enforcement and equal access to police protection services. Future work should assess the costs and benefits of different law enforcement practices that may be used to increase uniformity in officer behavior, including additional police training, monitoring protocols, mentorship programs, and targeted hiring and firing of officers.

1.8 Tables and Figures

Table 1.1: Summary Statistics

Table 1.1.A: Summary Statistics: Outcomes, Officers, and Complainants

	<i>Total Sample</i>		<i>Analysis Sample</i>	
	Mean	S.D.	Mean	S.D.
Total Observations	167,747		217,633	
Total Incidents	167,747		161,531	
Total Officers	2,904		1,851	
Outcomes				
Arrest	0.09	(0.29)	0.11	(0.31)
Suspect	0.17	(0.38)	0.19	(0.39)
Felony Arrest	0.03	(0.17)	0.04	(0.18)
Misdemeanor Arrest	0.07	(0.25)	0.08	(0.27)
Arrestee Black	0.03	(0.18)	0.04	(0.19)
Arrestee Hispanic	0.01	(0.11)	0.01	(0.12)
Arrestee White	0.01	(0.11)	0.01	(0.12)
Officer Characteristics				
Officer Arrest Rate	0.11	(0.06)	0.11	(0.05)
Two Responders	0.35	(0.48)	0.52	(0.50)
Black	0.25	(0.41)	0.24	(0.39)
Hispanic	0.20	(0.37)	0.21	(0.35)
White	0.50	(0.46)	0.50	(0.45)
Female	0.16	(0.33)	0.16	(0.32)
Age	39.31	(9.66)	38.37	(9.30)
Trainee	0.05	(0.19)	0.07	(0.20)
Sergeant	0.02	(0.13)	0.01	(0.10)
Salary (\$10,000s)	6.00	(1.08)	5.88	(1.02)
Years of Experience	12.70	(8.95)	11.82	(8.56)
Total Incidents	167.25	(107.87)	171.03	(103.62)
Complainant Characteristics				
Victim with Injury	0.08	(0.28)	0.10	(0.29)
Number of Complainants	1.56	(0.96)	1.59	(0.99)
Black	0.35	(0.48)	0.36	(0.48)
Hispanic	0.35	(0.48)	0.35	(0.48)
White	0.33	(0.47)	0.32	(0.47)
Female	0.49	(0.50)	0.50	(0.50)

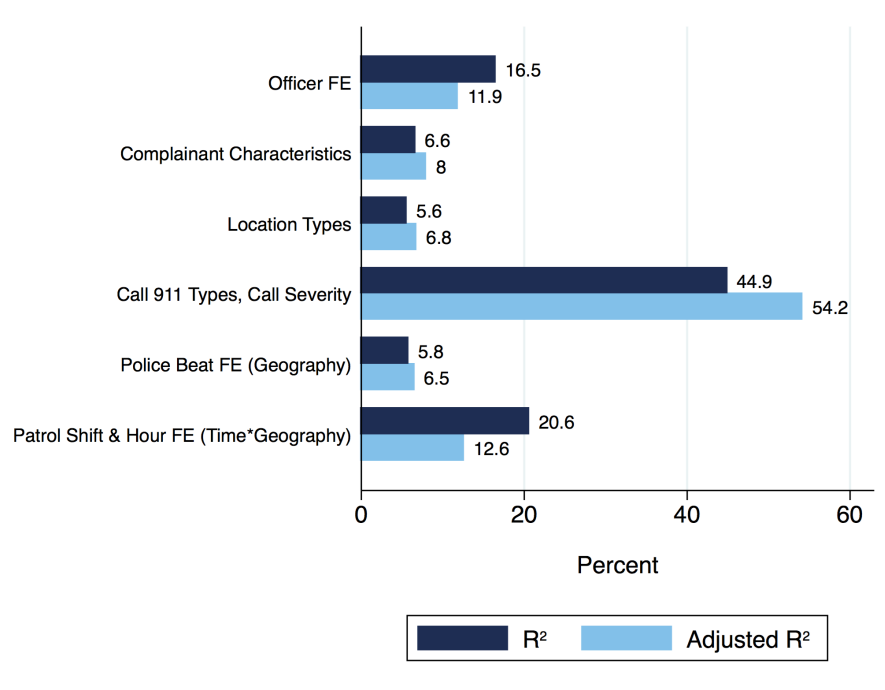
These tables display summary statistics for covariates used in analysis. The first column, “Total Sample”, consists of all incidents in the data, with each incident having only one record. The second column, “Analysis Sample”, summarizes the primary analysis sample and excludes records for police officers that respond to fewer than 25 incidents. Additionally, this sample duplicates incident responses with two responding officers so that the records for each responding officer may contribute to the estimation of Officer Fixed Effects. Complainant characteristics are shown for observations that have information on complainant demographics, which corresponds to 95% of the sample. Information for suspect outcomes is only available prior to 2017 and is summarized for the period available.

Table 1.1.B: Summary Statistics: Incident Urgency, Location Type, and Dispatch Code

	<i>Total Sample</i>		<i>Analysis Sample</i>	
	Mean	S.D.	Mean	S.D.
Total Observations	167,747		217,633	
Total Incidents	167,747		161,531	
Total Officers	2,904		1,851	
Call Urgency				
Time to Dispatch (Minutes)	24.81	(28.31)	24.09	(27.94)
Location Type				
Apartment	0.12	(0.33)	0.13	(0.33)
Residence Other	0.15	(0.36)	0.15	(0.36)
Bar/Club/Entertainment	0.03	(0.17)	0.03	(0.17)
Retail	0.06	(0.24)	0.07	(0.25)
Business Other	0.05	(0.22)	0.05	(0.22)
Govt/Health/School/Religion	0.01	(0.10)	0.01	(0.10)
Motor Vehicle	0.02	(0.15)	0.02	(0.15)
Parking Lot	0.25	(0.43)	0.24	(0.43)
Street	0.20	(0.40)	0.20	(0.40)
Outdoor Other	0.05	(0.22)	0.05	(0.22)
Other Location	0.05	(0.21)	0.05	(0.21)
Dispatch Code Type				
Criminal Assault	0.01	(0.12)	0.01	(0.12)
Armed Encounter/Active Shooter	0.02	(0.12)	0.02	(0.13)
Injured Person	0.01	(0.11)	0.01	(0.11)
Robbery	0.05	(0.22)	0.06	(0.24)
Burglary of Business	0.05	(0.22)	0.05	(0.22)
Burglary of Residence	0.12	(0.32)	0.11	(0.32)
Burglary of Motor Vehicle	0.19	(0.39)	0.17	(0.38)
Unauthorized Use of Motor Vehicle	0.06	(0.23)	0.05	(0.22)
Theft	0.08	(0.27)	0.07	(0.26)
Criminal Mischief	0.07	(0.26)	0.07	(0.25)
Major Disturbance	0.10	(0.30)	0.11	(0.32)
Major Accident	0.04	(0.20)	0.04	(0.21)
Minor Accident	0.10	(0.30)	0.09	(0.29)
Other - Serious	0.06	(0.24)	0.07	(0.26)
Other - Minor	0.04	(0.20)	0.04	(0.21)

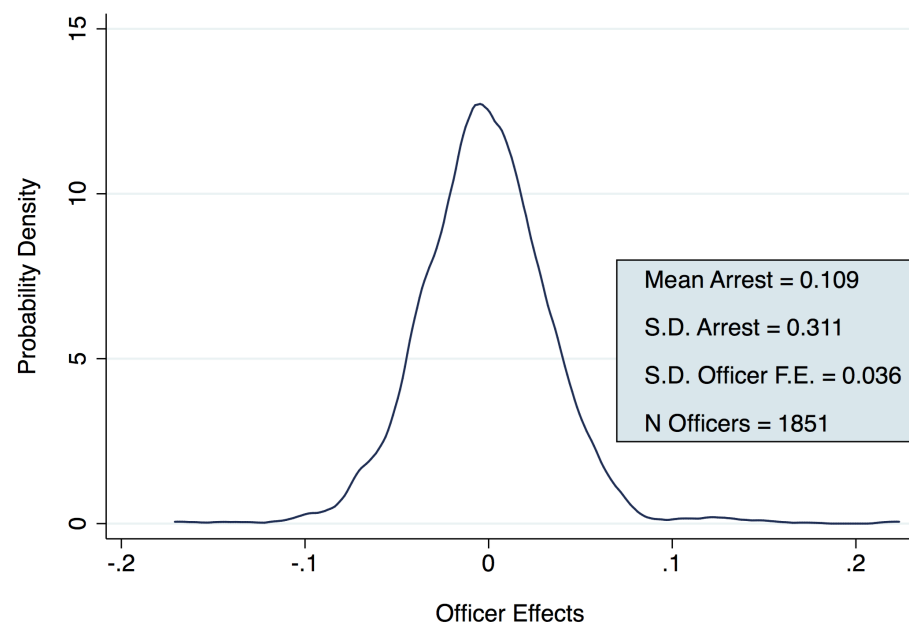
These tables display summary statistics for covariates used in analysis. The first column, “Total Sample”, consists of all incidents in the data, with each incident having only one record. The second column, “Analysis Sample”, summarizes the primary analysis sample and excludes records for police officers that respond to fewer than 25 incidents. Additionally, this sample duplicates incident responses with two responding officers so that the records for each responding officer may contribute to the estimation of Officer Fixed Effects. Call Urgency or Time to Dispatch (Minutes) is a variable that captures the priority level or severity of a given call at the time of dispatch.

Figure 1.1: Relative Proportion of Total R^2 , Components of Model



This figure shows the relative importance of explanatory variables in the estimation model. Each bar graph grouping represents the percent of total model R^2 (Adjusted R^2) accounted for by a grouping of variables. I estimate these percentages sequentially from the bottom bar to the top, with each percentage calculated as an additional contribution of R^2 to the total or: $(R^2_{currentbar} - R^2_{priorbar})/R^2_{total}$. The explanatory power of Officer FE is therefore the relative contribution of these variables after controlling for all other variables in the model.

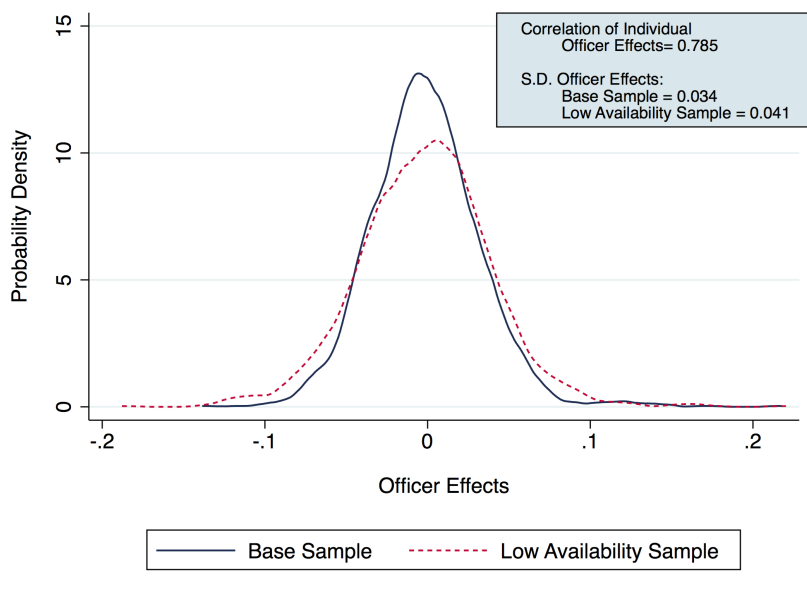
Figure 1.2: Distribution of Officer Effects, $\hat{\theta}_i$



This figure graphs the distribution of the Officer Effects, $\hat{\theta}_i$, measured in the primary arrest outcome model on the analysis sample. Each officer in the sample has at least 25 incident responses.

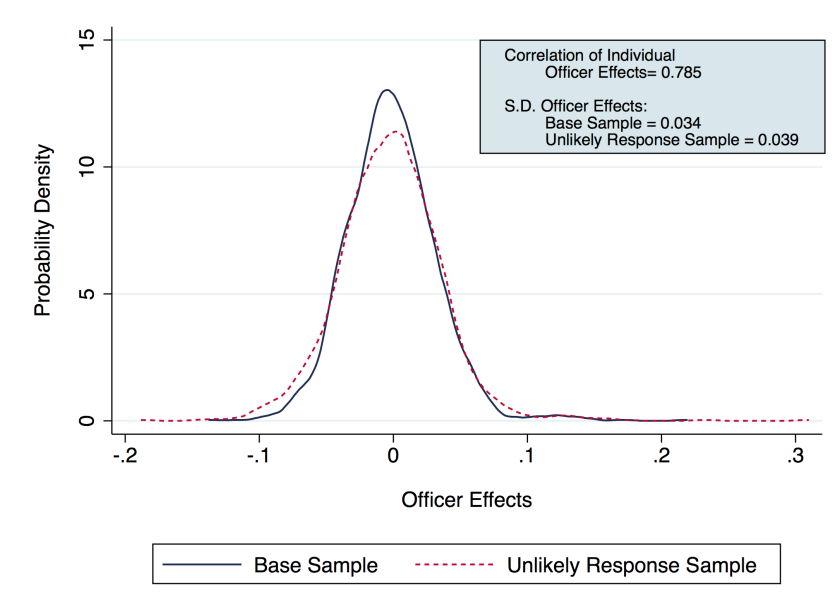
Figure 1.3: Tests of the Importance of Officer Sorting to Officer Effect Distribution

Figure 1.3.A: Officer Effects in Full Sample and Low Availability Sample



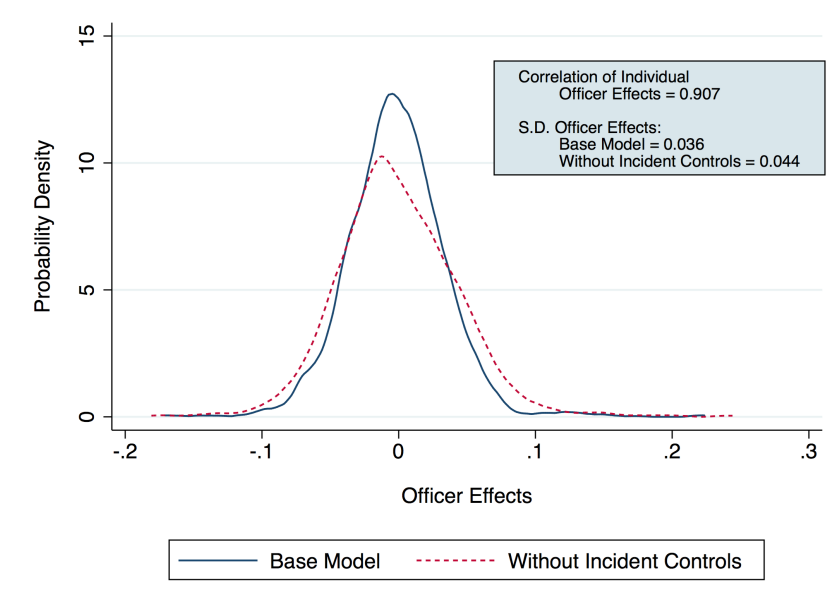
This graph compares the base model officer effects, $\hat{\theta}_i$, to officer effects, $\hat{\theta}'_i$, that are estimated from the “Low Availability” sample. The “Low Availability” sample is determined by taking the set of observations where more officers are unavailable because they are on other calls at the time a call is made (split at the median within patrol shift cells). The analysis is restricted to officers with at least 25 observations in the sub-sample. The corresponding base sample benchmark is estimated over the full set of responses for the same officer group.

Figure 1.3.B: Officer Effects in Full Sample and Unlikely Response Sample



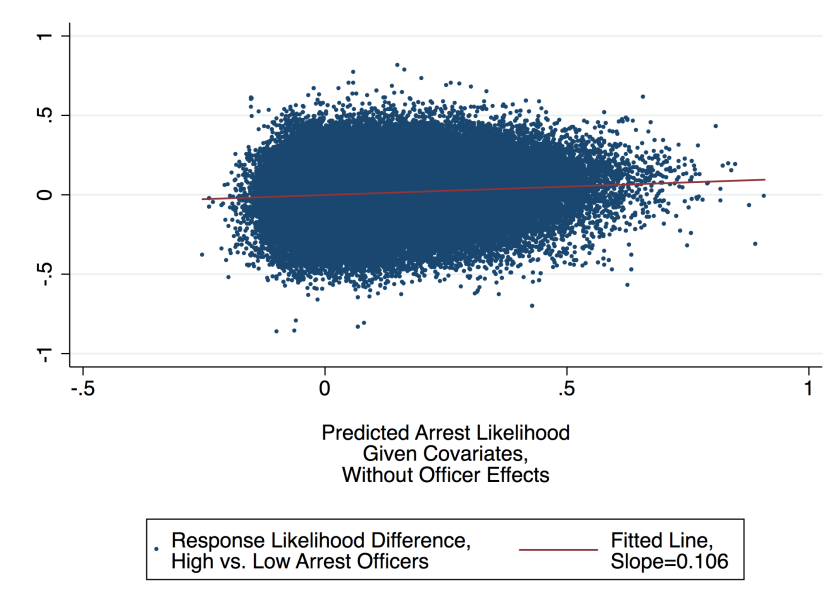
This graph compares the base model officer effects, $\hat{\theta}_i$, to officer effects, $\hat{\theta}_i''$, that are estimated from the “Unlikely Response” sample. The “Unlikely Response” is determined by taking a sub-sample of observations where officers have a lower predicted probability of responding to calls, where response probability is measured using a regression that predicts each officer’s response to an incident given observable characteristics (sample below median response likelihood for each officer). The analysis is restricted to officers with at least 25 observations in the sub-sample. The corresponding base sample benchmark is estimated over the full set of responses for the same officer group.

Figure 1.3.C: Officer Effects Measured with and without Incident Controls



This graph compares the base model officer effects, $\hat{\theta}_i$, to officer effects, $\hat{\theta}_i'''$, that are estimated from a model that does not include incident characteristics and police beat fixed effects, X_{kt} and ϕ_g . These incident characteristics are those that an officer may choose at the level of an incident response, as shifts and partners are determined prior to a call event.

Figure 1.3.D: Balance Test of Arrest Likelihood and Officer Response Likelihood



This graph examines whether high (low) officers are similarly likely to respond to incidents, given observable characteristics. Observable characteristics are aggregated as the predicted arrest probability of an incident, from a model without officer effects (X axis). Response likelihoods of high arrest officers (top 25% of officer effects) and low arrest officers (bottom 25% of officer effects) are determined by separately predicting responses for these officer groups using a model that excludes officer effects. The Y axis measures whether high arrest officers are more likely than low arrest officers to respond to a given incident using the difference between these two response likelihoods.

Table 1.2: Summary Results and Tests of Sorting, Comparison to Unlikely Response Sample

	Analysis Sample (1) Full Sample	Low Availability Sample (2) Corresponding Full Sample	(3) Low Availability Responses	Unlikely Response Sample (4) Corresponding Full Sample	(5) Unlikely Responses
Primary Results					
<i>Contribution of Officer Effects</i>					
Relative % of R-2 from Officer Effects	16.5%	15.9%	18.5%	15.8%	18.8%
Relative % of Adj. R-2 from Officer Effects	11.9%	11.6%	12.6%	11.5%	13.6%
<i>Distribution of Officer Effects</i>					
S.D. of Officer Effect	0.036	0.034	0.041	0.034	0.039
% Change: 1 S.D. Increase in Officer Effect	32.8%	31.2%	38.5%	31.3%	41.5%
Gap: 10th to 90th Percentile in Officer Effect	0.082	0.080	0.098	0.080	0.089
% Change: 10th to 90th Percentile in Officer Effect	75.9%	74.1%	92.5%	74.0%	94.5%
Auxiliary Results					
<i>Distribution of Officer Effects:</i>					
<i>Model without Incident Controls</i>					
S.D. of Officer Effect	0.044	0.042	0.048	0.042	0.047
% Change: 1 S.D. Increase in Officer Effect	40.8%	39.1%	45.1%	39.5%	49.6%
Gap: 10th to 90th Percentile in Officer Effect	0.105	0.102	0.117	0.102	0.110
% Change: 10th to 90th Percentile in Officer Effect	96.6%	94.4%	110.5%	94.7%	117.1%
<i>Correlation of Officer Effects</i>					
Full model and Model Without Incident Controls	0.907	0.901	0.910	0.906	0.923
Sub-sample and Corresponding Full Sample			0.785		0.785
Mean of Outcome	0.109	0.108	0.106	0.108	0.094
S.D. of Outcome	0.311	0.310	0.307	0.310	0.292
Total Officers	1,851	1,556	1,556	1,504	1,504
Total Observations	217,633	207,203	116,070	205,020	102,203

This table summarizes the main analysis arrest results and tests of officer sorting. The contribution of officer effects shows the relative proportion of R^2 due to including officer effects in the model. The "Correlation of Officer Effects" compares the base model officer effects, $\hat{\theta}_i$, to officer effects, $\hat{\theta}_i''$, that are estimated from a model that does not include incident characteristics and police beat fixed effects, X_{kt} and ϕ_g .

Table 1.3: Officer Effects and Officer Demographics

	(1)	(2)	(3)
<i>Outcome: Officer Effect</i>	Full Sample	Low Availability Responses	Unlikely Responses
Black	0.002 (0.004)	0.008 (0.005)	0.002 (0.005)
Hispanic	-0.002 (0.004)	0.003 (0.005)	-0.003 (0.005)
White	0.004 (0.004)	0.009+ (0.005)	0.004 (0.005)
Female	-0.004+ (0.002)	-0.003 (0.002)	-0.003 (0.002)
Age	-0.0004* (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0002)
Trainee	-0.010** -0.003	-0.003 (0.004)	-0.009* (0.004)
Sergeant	-0.005 (0.005)	-0.018 (0.011)	-0.015* (0.007)
Experience	0.0019*** (0.0004)	0.0016** (0.0005)	0.0023*** (0.0005)
Experience^2	-0.00004*** (0.00001)	-0.00003* (0.00001)	-0.00005*** (0.00001)
Observations	1,832	1,551	1,500
R-squared	0.041	0.024	0.042
Fixed Effect Mean	-0.001	-0.002	-0.002
Fixed Effect S.D.	0.036	0.041	0.039
Outcome Mean	0.109	0.106	0.094
Outcome S.D.	0.311	0.307	0.292

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table shows regression results of officer effects measured using the arrest outcome, $\hat{\theta}_i$, regressed on fixed officer characteristics, at the officer level. Robust standard errors are in parentheses. The analysis in columns (2) and (3) is restricted to officers with at least 25 observations in each sub-sample. Other race officers are the omitted race category. Officers without demographic information are excluded from the regressions.

Figure 1.4: Bootstrap Benchmark Test Distribution

Figure 1.4.A: S.D. of Officer Effect Distribution

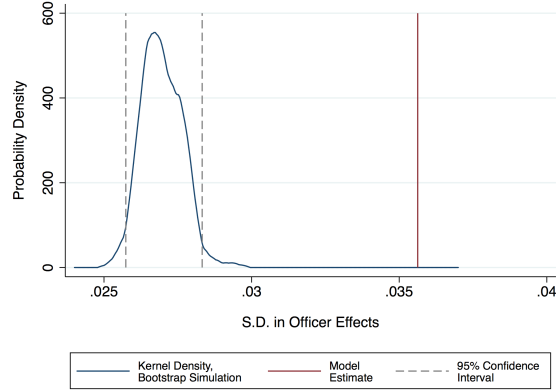
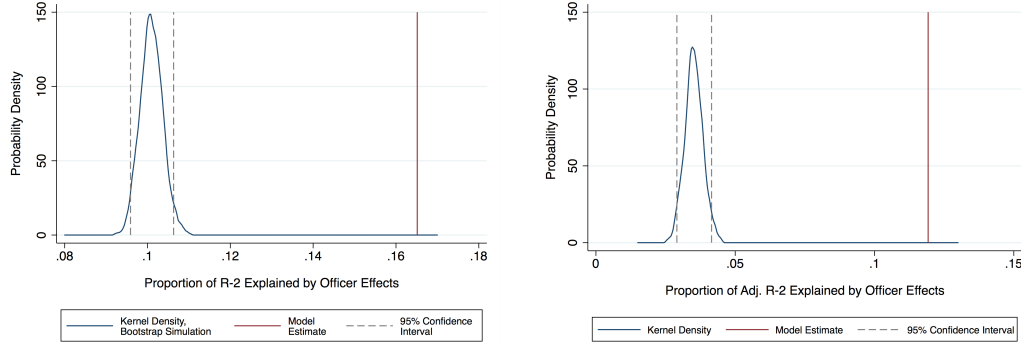


Figure 1.4.B: Proportion of R^2 (left) and Adjusted R^2 (right) Explained by Officer Effects



Each graph shows the residual bootstrap test distribution for three primary statistics, S.D. of the Officer Effect distribution, as well as the relative proportion of R^2 and Adj. R^2 explained by officer fixed effects. Each bootstrap iteration is obtained as follows: (1) Residuals and predicted outcomes are obtained from a first stage model that does not include Officer FE (under the null hypothesis that these variables are jointly zero), (2) Estimated residuals are assigned a wild bootstrap weight of $w \in \{1, -1\}$ that is constant within shift clusters δ_{gt} , with equal probability for each shift group, and these residuals are added to the predicted outcomes from (1), (3) Using these simulated outcome variables, the full model, including Officer FE, is estimated to obtain each statistic of interest. Post-estimation Empirical Bayes adjustments are made to the estimates after each iteration. Each test is based on 250 bootstrap replications.

Table 1.4: Robustness Specification Tests

Table 1.4.A: Summary Results, Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Base Model	Reporting Area Fixed Effects	Sector*Month Fixed Effects	Individual Shift Effects	Add Full Set of Dispatch & Location Codes	First Stage Fixed Effects: >100 N	First Stage Fixed Effects: Weighted by N	First Stage Fixed Effects: Unadjusted
% of R-2 from Officer Effects	16.5%	15.8%	16.5%	10.4%	11.2%	15.4%	16.7%	16.7%
% of Adj. R-2 from Officer Effects	11.9%	11.5%	12.1%	9.1%	7.1%	11.5%	12.0%	12.0%
S.D. of Officer Effect	0.036	0.036	0.037	0.044	0.033	0.039	0.041	0.050
% Change: 1 S.D. Increase in Officer Effect	32.8%	33.7%	33.6%	40.8%	30.8%	38.0%	37.9%	46.1%
Gap: 10th to 90th Percentile in Officer Effect	0.082	0.085	0.085	0.101	0.077	0.086	0.092	0.107
% Change: 10th to 90th Percentile in Officer Effect	75.9%	78.1%	78.0%	92.9%	71.2%	82.9%	84.3%	98.4%
Correlation to Arrest Officer Effect		0.992	0.995	0.949	0.970	0.976	0.982	0.971
Mean of Outcome	0.109	0.108	0.109	0.109	0.109	0.103	0.109	0.109
S.D. of Outcome	0.311	0.311	0.311	0.311	0.311	0.305	0.312	0.311
Total Officers	1,851	1,851	1,851	1,851	1,851	942	1,851	1,851
Total Observations	217,633	215,539	217,633	217,256	205,829	163,433	205,829	217,633

This table summarizes the analysis results across robustness specifications. Column (1) replicates the results from the primary specification. In column (2), police beat FE, ϕ_g , are replaced with reporting area FE, a finer geographic unit. In column (3), police beat FE, ϕ_g , are replaced with Sector*Month FE, which interact the 35 geographic police sectors with month indicators. Column (4) restricts the sample to police responses for officers with over 100 incident responses in the sample. Column (5) inserts the full set of 117 dispatch codes and 34 location type codes available in the data, rather than the broader 15 dispatch codes and 11 location type codes used in the preferred specification. Column (6) reports the dispersion in unadjusted officer fixed effects from the first stage, using a sample restricted to officers with more than 100 observations. Column (7) calculates the dispersion metric as a standard deviation of unadjusted officer fixed effects from the first stage that is weighted by the number of observations for each officer. In column (8) the correlation to base model officer effects is weighted by the number of observations for each officer. Column (8) reports dispersion in the unadjusted officer fixed effect estimates.

Table 1.4.B: Correlation of Officer Fixed Effects, Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Base Model	Reporting Area Fixed Effects	Sector*Month Fixed Effects	Add Individual Shift Effects	Add Full Set of Dispatch & Location Codes	First Stage Fixed Effects: >100 N	First Stage Fixed Effects: Unadjusted
(1) Base Model	1.00						
(2) Add Reporting Area Fixed Effects	0.99	1.00					
(3) Add Sector*Month Fixed Effects	1.00	0.99	1.00				
(4) Add Individual Shift Effects	0.95	0.94	0.95	1.00			
(5) Add Full Set of Dispatch Codes	0.97	0.96	0.96	0.92	1.00		
(6) First Stage Fixed Effects: >100 N	0.98	0.97	0.97	0.95	0.95	1.00	
(7) First Stage Fixed Effects: Unadjusted	0.97	0.97	0.97	0.95	0.94	0.98	1.00

This table summarizes the analysis results across robustness specifications. Column (1) replicates the results from the primary specification. In column (2), police beat FE, ϕ_g , are replaced with reporting area FE, a finer geographic unit. In column (3), police beat FE, ϕ_g , are replaced with Sector*Month FE, which interact the 35 geographic police sectors with month indicators. Column (4) restricts the sample to police responses for officers with over 100 incident responses in the sample. Column (5) inserts the full set of 117 dispatch codes and 34 location type codes available in the data, rather than the broader 15 dispatch codes and 11 location type codes used in the preferred specification. Column (6) reports the dispersion in unadjusted officer fixed effects from the first stage, using a sample restricted to officers with more than 100 observations. Column (7) calculates the dispersion metric as a standard deviation of unadjusted officer fixed effects from the first stage that is weighted by the number of observations for each officer. In column (8) the correlation to base model officer effects is weighted by the number of observations for each officer. Column (8) reports dispersion in the unadjusted officer fixed effect estimates. The specifications in column (7) in Panel (A) are removed from Panel (B) because it involves weighted data. Correlations shown in the bottom panel are calculated for overlapping observations across different specifications.

Table 1.5: Racial Bias Test, Officer and Arrestee Race

Table 1.5.A: Racial Bias Test, Full Sample

		(1)	(2)	(3)	(4)
<i>Outcome=Officer Effects</i>		Arrest	Arrest Black	Arrest Hispanic	Arrest White
A. Full Sample					
Black Officer		-0.001 (0.005)	0.0004 (0.002)	0.0002 (0.001)	-0.0005 (0.001)
Hispanic Officer		-0.004 (0.005)	-0.0001 (0.002)	-0.002 (0.001)	-0.001 (0.001)
White Officer		0.003 (0.004)	0.002 (0.002)	0.0003 (0.001)	-0.0004 (0.001)
Black=Hispanic:	F-Test	0.94	0.10	5.15	0.00
	P-Value	0.33	0.75	0.02	0.98
Black=White:	F-Test	2.84	2.64	0.02	0.01
	P-Value	0.09	0.11	0.88	0.91
Hispanic=White:	F-Test	7.26	4.37	6.49	0.02
	P-Value	0.01	0.04	0.01	0.89
Observations		1,613	1,613	1,613	1,613
Arrest Mean		0.091	0.034	0.013	0.012

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table shows regressions of officer specific effects on officer demographics, where officer effects are derived from four different first stage outcomes, general arrests, and whether the arrestee was Black, White or Hispanic. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Robust standard errors are in parentheses. The full sample, Panel (A), is restricted to observations where responding officers have a single race, and each officer has more than 25 observations within this restriction.

Table 1.5.B: Racial Bias Test, “Low Availability” and “Unlikely Response” Samples

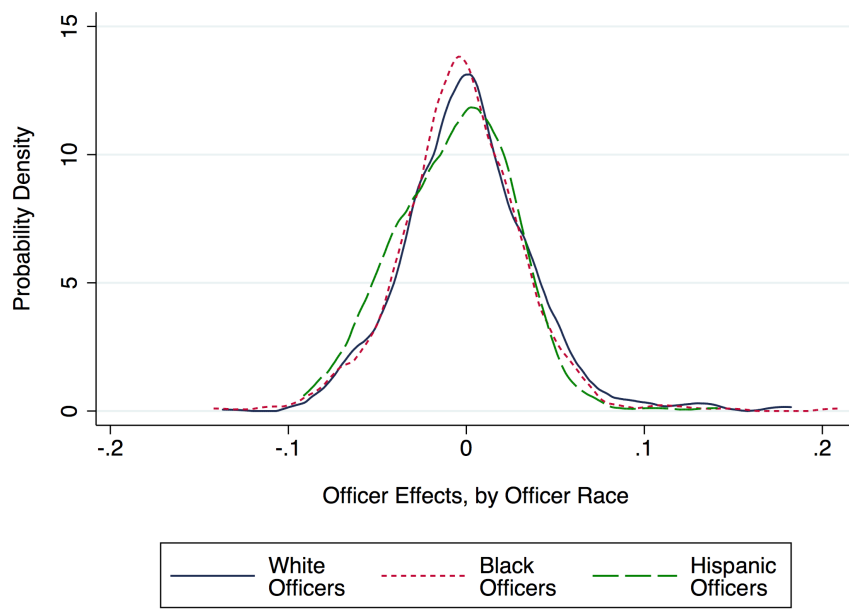
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Outcome=Officer Effects</i>		Arrest	Arrest Black	Arrest Hispanic	Arrest White	Arrest	Arrest Black	Arrest Hispanic	Arrest White
B. "Low Availability" Sample									
Black Officer		-0.003 (0.006)	-0.001 (0.003)	-0.001 (0.002)	0.004* (0.002)	-0.002 (0.005)	-0.0001 (0.003)	-0.001 (0.002)	0.001 (0.001)
Hispanic Officer		-0.006 (0.006)	-0.005 (0.003)	-0.003 (0.002)	0.004+ (0.002)	-0.004 (0.005)	-0.002 (0.003)	-0.002 (0.002)	0.001 (0.002)
White Officer		0.002 (0.006)	-0.0003 (0.003)	0.0003 (0.002)	0.003 (0.002)	0.002 (0.005)	0.001 (0.003)	-0.0002 (0.002)	0.001 (0.001)
Black=Hispanic:	F-Test P-Value	0.70 0.40	2.75 0.10	1.29 0.26	0.19 0.66	0.32 0.57	1.37 0.24	0.52 0.47	0.28 0.59
Black=White:	F-Test P-Value	2.97 0.09	0.49 0.48	2.81 0.09	1.69 0.19	2.13 0.14	0.57 0.45	1.15 0.29	0.30 0.59
Hispanic=White:	F-Test P-Value	5.40 0.02	6.53 0.01	5.67 0.02	0.42 0.52	4.15 0.04	4.74 0.03	2.71 0.10	0.01 0.93
Observations		1,254	1,254	1,254	1,254	1,308	1,308	1,308	1,308
Arrest Mean		0.089	0.032	0.013	0.013	0.077	0.028	0.011	0.010
C. "Unlikely Response" Sample									
Black Officer									
Hispanic Officer									
White Officer									
Black=Hispanic:	F-Test P-Value								
Black=White:	F-Test P-Value								
Hispanic=White:	F-Test P-Value								
Observations									
Arrest Mean									

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table shows regressions of officer specific effects on officer demographics, where officer effects are derived from four different first stage outcomes, general arrests, and whether the arrestee was Black, White or Hispanic. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Robust standard errors are in parentheses. Panels (B) and (C) represent the overlap of Panel (A) with the “Low Availability” and “Unlikely Response” samples, where the sample is restricted to officers with at least 25 observations within each subset. The interaction of arrestee outcome race and officer race through the regression coefficients represents a test of officer and arrestee race interaction effects. F-Tests measure whether officers of different races are more likely to make arrests of individuals of different races.

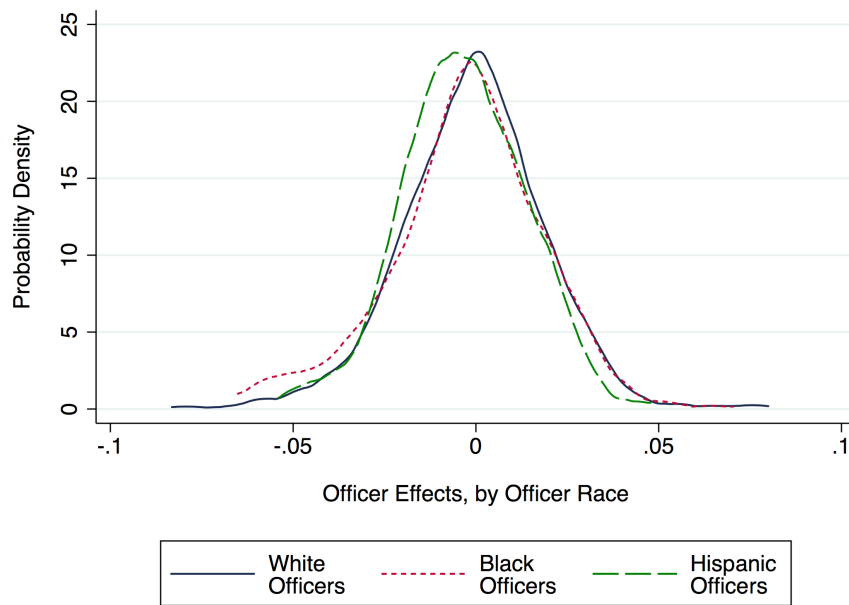
Figure 1.5: Distribution of Officer Effects across Officer Race

Figure 1.5.A: Officer Effects by Officer Race, Arrest Outcome



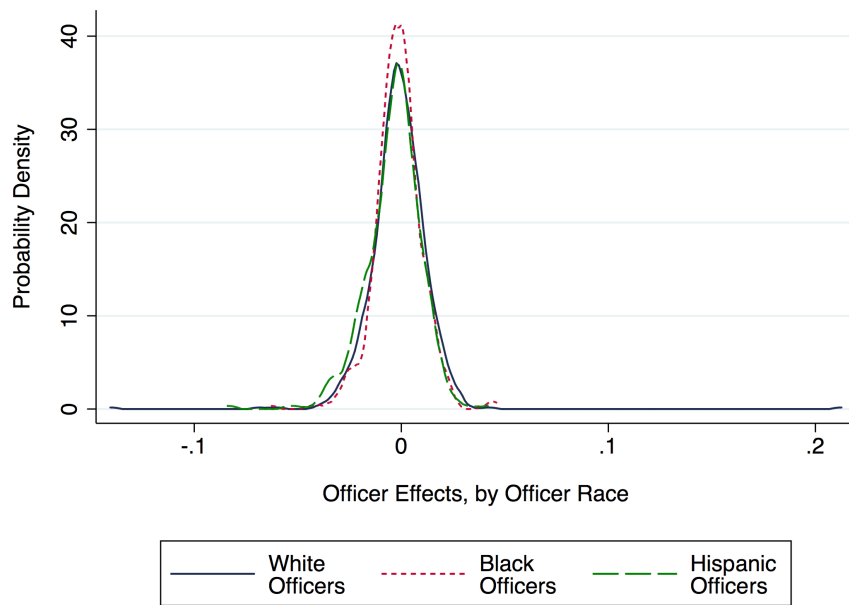
These graphs the distribution of officer effects by officer race for different arrest demographic outcomes. The officer effects correspond to the outcomes used in Table 1.5. These graphs are shown for the full racial bias test sample, which restricts to observations where the responding officers have a single race and each officer within this restriction has at least 25 observations. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Each graph shows the density overlap of officer effects for the different arrest outcomes.

Figure 1.5.B: Officer Effects by Officer Race, Arrestee Black Outcome



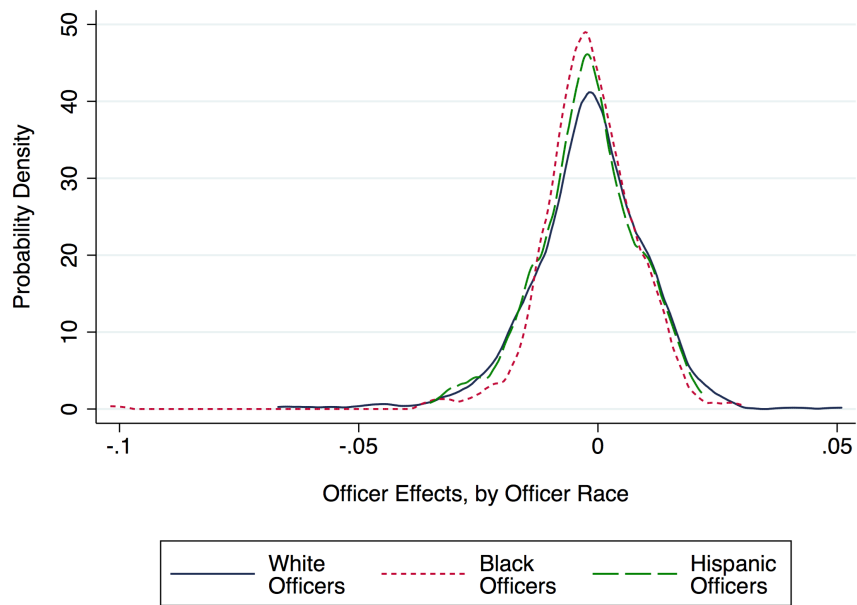
These graphs show the distribution of officer effects by officer race for different arrest demographic outcomes. The officer effects correspond to the outcomes used in Table 1.5. These graphs are shown for the full racial bias test sample, which restricts to observations where the responding officers have a single race and each officer within this restriction has at least 25 observations. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Each graph shows the density overlap of officer effects for the different arrest outcomes.

Figure 1.5.C: Officer Effects by Officer Race, Arrestee Hispanic Outcome



These graphs show the distribution of officer effects by officer race for different arrest demographic outcomes. The officer effects correspond to the outcomes used in Table 1.5. These graphs are shown for the full racial bias test sample, which restricts to observations where the responding officers have a single race and each officer within this restriction has at least 25 observations. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Each graph shows the density overlap of officer effects for the different arrest outcomes.

Figure 1.5.D: Officer Effects by Officer Race, Arrestee White Outcome



These graphs show the distribution of officer effects by officer race for different arrest demographic outcomes. The officer effects correspond to the outcomes used in Table 1.5. These graphs are shown for the full racial bias test sample, which restricts to observations where the responding officers have a single race and each officer within this restriction has at least 25 observations. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Each graph shows the density overlap of officer effects for the different arrest outcomes.

Chapter 2

Safety in Police Numbers: Evidence of Police Effectiveness from COPS Grant Applications

Understanding the impact of police on crime is critical to designing policies that maximize safety. In this paper, I use a novel estimation approach to measure the impact of police hiring, which exploits variation in federal Community Oriented Policing Services (COPS) hiring grants, while also controlling for the endogenous decisions of police departments to apply for these grants. Using data from nearly 7,000 U.S. municipalities, I find that a 10% increase in police employment rates reduces violent crime rates by 13% and property crime rates by 7%. The results also provide suggestive evidence that law enforcement leaders are forward-looking.

2.1 Introduction

Policymakers, among others, are increasingly focused on the need for criminal justice reform. U.S. incarceration rates are over 300% larger than the world average and rates of police employment are 35% less than the world

average [51, 99]. At the same time, levels of policing per capita have remained relatively constant in the United States, declining by approximately 2% between 1995 and 2014.¹ In recent years, pressures on local government budgets have put additional strain on police departments [21]. The lower emphasis on investment in policing relative to incarceration does not make sense in light of substantial research evidence that incarceration rates are ineffective at current levels.² But at what cost? What are the benefits of additional police? This paper contributes to this important policy debate by evaluating the impact of police force expansions on reported crime rates.

Measuring the causal impact of police on crime is difficult because police hiring is likely a cause as well as a consequence of changing criminal activity. It is not surprising that police districts with high levels of crime also have large police departments or that districts with a spike in crime may expand resources for law enforcement. In order to measure the causal impact of police, it is necessary to use experimental or quasi-experimental sources of variation in police staffing that are not driven by crime. The existing economics literature studying the effect of police has leveraged a variety of sources of quasi-experimental variation, including variation in local election cycles, terrorist attacks, state tax rates, and mass lay-offs [e.g. 25, 29, 31, 64, 70, 71, 72, 76].

¹This change is calculated using data on the total number of sworn police from national FBI Uniform Crime Report (UCR) data, adjusted by national population estimates from the U.S. Census.

²A large literature finds that there are decreasing marginal returns to incarceration rates in terms of crime reduction. Because U.S. levels are so high, investments in incarceration are likely ineffective. See for example [30, 15, 60, 80].

Recent work by [15] also finds that measurement error in police and crime data sources may also cause bias in estimates of the causal impact of police. A number of controlled experiments conducted by criminologists within particular cities have also found that increasing policing presence in targeted crime "hot spots" reduces crime in these locations [e.g. 9, 101]. My estimation strategy builds upon this research by providing both a new estimate of the causal impact of police hiring and by developing empirical insights about how law enforcement decisions to invest in hiring are related to changes in crime.

In this paper, I use variation in Community Oriented Police Services (COPS) hiring grant awards and grant applications to measure the impact of police force expansions on crime. Since its establishment in 1994 under the Violent Crime Control Act (VCCA), the COPS Office at the Department of Justice (DOJ) has distributed over \$12 billion in total grants to fund police hiring, invest in policing technology, support police officers in public schools, advance community policing, supplement law enforcement funding in small towns and tribal communities, and target specific crimes. Approximately \$7 billion of COPS grant funding has been devoted to police hiring grants, designating support for over 80,000 eligible sworn police officer positions. This paper increases our understanding of this program in recent decades, focusing on the effectiveness of grants administered between 2000-2014.

Importantly, because the decision to apply for police hiring grants is not random, I compare crime outcomes within cities in years when they

receive hiring funding to outcomes in years when their grant applications are rejected, as well as to baseline years when cities did not seek funding. Controlling for grant applications is an important feature of my strategy because police departments seek federal hiring funds when they are interested in expanding their police forces. This interest depends on the needs and goals of departments, which are in turn related to fluctuations in local crime conditions and expectations about future crime. In practice, I measure the impact of police on crime using an instrumental variables design, instrumenting for municipal police presence using COPS grant acceptances conditional on grant application decisions. Because there are large fluctuations in total funding for the COPS program office, there is substantial variation in the number of possible grants that can be awarded in each year. This program variation creates quasi-randomness in the probability of a grant acceptance in a given year within a given city and is the foundation of my identification strategy.

In concurrent work, [77] and [20] study COPS grants awarded since 2009 using regression discontinuity designs. While similar to this paper in spirit, there are several notable differences between our papers. First, I emphasize panel data comparisons *within* police districts over time across years when departments receive grants versus years in which they do not receive grants, over a longer period and broader set of cities than has been used in other work. The sample in this paper covers approximately 10 times the number of cities examined in [77] and 7 times the number of cities as in [20]. Second, because I use information on multiple years of the COPS grant pro-

gram, I can control for overlapping COPS hiring grant cycles as well as COPS grant applications and acceptances for other police functions (such as technology, school grants, and community partnership grants). Third, by directly including application controls in my model, I am able to empirically observe how application decisions are linked to subsequent changes in crime, a new contribution to the literature. The results across these papers are comparable, particularly when comparing the elasticities from the robustness checks with similar sample restrictions.³ In contrast, prior papers on the COPS hiring program have treated the timing of grant receipt as quasi-random, measuring crime outcomes within a city before and after the city received a grant to hire police officers [42, 34, 106].

In my primary specification, I find that a 10% increase in police employment rates reduces violent crime rates by 13% and property crime rates by 7%, estimates that overlap with the upper range of effects in the existing literature. In addition to the contribution of the study design, the analysis also uses a larger and more representative national data set than has been used in prior work, measuring effects in a sample of nearly 7,000 municipalities across the country between 2000 and 2014. The model estimates are comparable to prior work studying the impact of police, which have found crime elasticities

³Both [77] and [20] focus on grants administered in 2009. [77] uses information on COPS grant applications to define his sample comparison group and does not directly control for applications. The sample restrictions used in these papers are comparable to robustness checks in this paper that limit the sample to police departments that applied for or received a grant between 2000-2014 (Specification 9 and 10 in Table 2.6).

ranging from -0.3 to greater than -1, depending on the setting and study design.

Interestingly, I also find that police departments that apply for hiring grants also experience increases in crime rates in later years. Because of this relationship, estimates of grant acceptance that do not control for application decisions will be smaller than the true crime reducing effect of police. The fact that police departments experience increases in crime after they apply for hiring grants suggests that police departments may be forward-looking when they make investment decisions. These patterns support the notion that local police departments may be most knowledgeable of their own resource needs, and that when they seek federal funding but are denied assistance, they operate less efficiently.

I also explore the mechanisms of the crime reducing impact of police hiring. I find that solved or "cleared" crimes decline at a slower rate than reported crimes when police forces expand, providing evidence that investments in police hiring increase the rate of crime-solving. At the same time, police hiring does not increase the total level of arrests or "cleared" crimes for serious crime categories, suggesting that police hiring reduces crime through deterrence rather than incapacitation. Examining a broader set of arrest categories, I also find that when police departments expand, arrest rates for narcotics crimes decline, while arrest rates increase for marijuana crimes and driving under the influence of alcohol (DUI). Using data on arrests for dif-

ferent demographic subgroups, I show that the changes in drug crime resulting from police hiring are concentrated among Black males, while the increase in DUI arrests is concentrated among White males.

This paper is organized as follows: Section 2.2 provides details of the federal COPS hiring grant program and institutional context for the estimation strategy. Section 3.3 describes the empirical model and the data sources used in estimation. Section 2.4 discusses the results, outlining main findings as well as robustness tests and analysis of potential mechanisms. Section 2.5 concludes.

2.2 COPS Hiring Grant Program

In 1994, the VCCA authorized the Attorney General to distribute federal funds to local law enforcement through the COPS Office at the Department of Justice, with the goal of putting 100,000 new police officers on the street by the year 2000. The COPS Office provides grants to law enforcement to support a number of functions, including technology improvements, public school safety, and community policing programs; however, hiring grants represent over 50% of funds distributed. Hiring grants have retained the same basic features over time, including provisions that grant funding is not used to supplant local funds for existing officers, that local law enforcement must match a portion of the funding for each new officer hired, and that each grant

is funded for period of three years.⁴

Early evaluations of COPS hiring grants by academic criminologists and the government found evidence that the grants were often used to supplant local funding for existing police, rather than hire new officers [41, 88, 106]. Given concerns about the effectiveness of the grants, the Bush Administration sharply decreased funding for the COPS office between 2000 and 2008, reaching a low of approximately \$210 million in new grants (of all types) distributed in 2007, with no new hiring grants distributed in 2006 or 2007.⁵ After the financial crisis in 2008, the Obama Administration advocated for increased funding for the COPS Office, as the Administration viewed COPS grants as a tool to support local law enforcement agencies with stretched resources during the recession. In 2009, the American Reinvestment and Recovery Act (ARRA) provided an additional \$1 billion in funds to the COPS Office to help address the personnel needs of state, local, and tribal police departments [91].

Descriptive statistics of COPS grants over time reflect the changing political interest in this program. The top two graphs in Figure 3.1 show

⁴The institutional details described in this section were obtained through the COPS website (<http://www.cops.usdoj.gov/>) as well as direct communications with staff at the COPS Office.

⁵In Figure 3.1, this dollar figure is lower, at less than \$200 million, because information in the figures only include data for the restricted analysis sample used in this study. The most important sample restriction is that I limit attention to municipal police agencies, rather than including the universe of different police department types (county police, sheriff's departments, state police, university police, etc.).

that grants and grant funding declined between 2000 and 2008, and then shot up again in 2009 with the increase in ARRA stimulus funding. The bottom left graph in Figure 3.1 shows that prior to 2000, nearly all hiring grants were accepted, but when funding was cut for the COPS office during the Bush Administration, grants became more competitive.⁶ The gap between applications and acceptances increased in 2009, as the demand for federal assistance during the downturn outstripped the newly authorized funding. Because the research design in this paper hinges on controlling for the timing of endogenous decisions to apply for hiring grants, I restrict the analysis to the time period between 2000-2014 when there is variation in hiring grant applications and acceptances. The bottom right graph shows the actual variation used in this study, where acceptances are defined during the full period that each grant is administered rather than just the first year a grant is awarded, and applications are similarly defined for the corresponding full treatment period when a grant is intended to be received.

Prior to 2009, COPS hiring grants provided up to 75% of the salary and benefit cost per police officer up to a max benefit of \$75,000 per officer over three years. During this period, police departments submitted short narrative

⁶Literature examining earlier years of this program could not exploit variation in COPS grant applications because nearly all hiring grants were approved by the COPS office prior to 2000. [34] use data from approximately 2,000 cities to measure the impact of grants awarded between 1994 and 2001. The authors conduct limited tests that show that total amount awarded to each department is not correlated with pre-program crime rates, but these tests are aggregated across years and are not referenced to the actual point in time when a grant is received.

applications for hiring grants, and the selection process for hiring grant recipients was based on the subjective determination of the COPS Office. These descriptions varied widely, but typically included a brief description of how prospective grant funds would be used. Likewise, the training, monitoring and auditing of grantees was not standardized in this period. Beginning in 2009, the COPS Office changed the terms of its grants to respond to greater personnel needs of police departments during recession, providing 100% of the salary and benefit cost of a police officer up to a max benefit of \$125,000 per officer over three years, for up to 5% of the current sworn police force.⁷ At the same time, the COPS Office formalized grant processes and increased its monitoring efforts of grant recipients.⁸

The new grants also had more extensive application procedures, including questions on district crime, financial needs, and community policing programs and partnerships.⁹ These applications were scored by considering

⁷Departments have also been limited to a maximum number of hires, and these hiring ceilings varied moderately year to year for police departments of different sizes, from 15-50 total officers.

⁸In this period, recipients were required to submit quarterly progress reports to the COPS Office and 10% of grantees were identified by a risk algorithm for more extensive monitoring. Relative to the number of grants that are monitored, only a tiny fraction are formally audited, with only 211 audits of COPS grantees (for all program types) between 2000-2014 (<https://oig.justice.gov/reports/cops-ext.htm>).

⁹Applicants must provide a strategic plan of community policing activities the department will enact during the grant period. For fiscal need, police departments provide information on their police force and hiring request and their budget, as well as district information on local tax revenue, the percentage of families in poverty, the unemployment rate, the foreclosure rate, and details on major financial or budgetary events in the district. Lastly, applicants are asked to provide counts of the 7 Index I crimes for the prior three years (<http://www.cops.usdoj.gov/Default.asp?Item=2819>).

the fiscal needs and community policing proposals in each district, while also placed a lower weight on local crime conditions. The application weight on crime varied from 20-35% of the total score between 2009 and 2014, while the weight on fiscal need ranged from 30-75% and the weight on community policing ranged from 15-50% of the total score. After considering application scores, the COPS office has faced the additional allocation constraint that at least 0.5% of grant program funds must be distributed to each state and half of all funds must go to police departments serving districts with fewer than 150,000 residents.

The fact that crime conditions were included in the hiring grant application process is an important consideration for the identification strategy in this paper, which presumes that the timing of a grant acceptance is not based on changes in crime rates within a district. From a grant design perspective, it is important to note that crime was a minor factor in grant distribution decisions that was weighed alongside a number of other dimensions and allocation restrictions. Given the importance of this consideration, I formally test this identification assumption empirically in Section [2.4.3](#) below.

2.3 Empirical Model and Data Sources

A primary obstacle to measuring the impact of crime control policy is simultaneity.¹⁰ While police may have a causal impact on reducing crime, districts with higher levels of crime may also respond by hiring more police. If the true effect of police is to reduce crime, then simultaneity in the model will cause the key coefficient on police to be upwardly biased toward zero.

This paper employs an instrumental variables design to break this feedback loop, using an instrument that increases police force size but is not a function of crime. The empirical design uses the discontinuous increase in funding for police from federal COPS hiring grants as an instrument for police force size, as has been used in other prior research [e.g. 34]. However, the design in this paper controls for the elective decision to apply for a COPS grant, measuring the effect of receiving a grant conditional on this application decision. By controlling for grant application decisions, this instrument is not a function of the decision of a police department to apply for a grant but is instead only a function of the federal government’s decisions about how to allocate these grants.

¹⁰As noted by [15], measurement error in policing data may also introduce bias in this relationship.

The empirical model is as follows:

$$\begin{aligned}
Police_{it} = & \alpha_0 + \alpha_1 AcceptHiring_{it} + \alpha_2 ApplyHiring_{it} \\
& + \alpha_3 ApplyOtherGrants_{it} + \alpha_4 AcceptOtherGrants_{it} \\
& + \gamma X_{it} + \delta_{ct} + \phi_i + u_{it}
\end{aligned} \tag{2.1}$$

$$\begin{aligned}
Crime_{it} = & \beta_0 + \beta_1 Police_{it} + \beta_2 ApplyHiring_{it} \\
& + \beta_3 ApplyOtherGrants_{it} + \beta_4 AcceptOtherGrants_{it} \\
& + \tilde{\gamma} X_{it} + \tilde{\delta}_{ct} + \tilde{\phi}_i + \varepsilon_{it}
\end{aligned} \tag{2.2}$$

where, $Crime_{it}$ and $Police_{it}$ are the crime rate and police rate for city i in year t , or the number of reported crimes or police officers per 10,000 residents.¹¹ The primary outcomes considered in this paper are the 7 Index I crimes, which include 4 violent crime categories, murder, rape, robbery, and aggravated assault, and 3 property crime categories, burglary, larceny, and vehicle theft. X_{it} is a vector of demographic covariates that includes municipal district population, proportion male, the racial distribution of the population (proportion White, Black, and Hispanic), the age distribution of the population (proportion under 14, between 15 and 24, and between 25 and 39), the unemployment rate, and average earnings for individuals employed in the municipality.

¹¹In this study, I use a liberal population restriction of cities with more than 1,000 residents. In my sample, several of the crime outcomes frequently take values of 0, rendering a log-log model infeasible without systematically restricting the sample to higher crime locations and years. Instead, I construct crime outcomes and police variables as rates per capita, which improves the model's comparison across cities of different sizes. While several papers scale crime rates by 100,000 residents, I adjust rates per 10,000 residents given the large number of smaller cities in the sample.

δ_{ct} are year by population group fixed effects, that allow the model to separately control for differing crime time trends in cities of different sizes. City population groups are split into cities with 1,000 to 1,999 residents, 2,000 to 4,999 residents, 5,000 to 9,999 residents, 10,000 to 24,999 residents, 25,000 to 49,999 residents, 50,000 to 99,999 residents and 100,000 residents or more.¹² In Section 3.4.4, I examine alternative controls for differing time trends across different district types, including robustness checks with state by year fixed effects and grant application cohort by year fixed effects.

The crime and police data are drawn from the Federal Bureau of Investigation (FBI) Uniform Crime Report (UCR) data, while the demographic data are compiled from the U.S. Census and the Bureau of Labor Statistics (BLS). Police agencies report their crime and police information to the UCR program voluntarily, and while the FBI conducts some auditing and monitoring of submissions, a number of reporting errors likely persist in the data. Because reporting errors may be more common for smaller police departments, prior literature has restricted analysis to larger cities, typically to cities over 10,000 residents or more. To address these issues and expand my sample to municipalities with 1,000 residents or more, the crime and police data are cleaned using a regression algorithm to identify outliers and replace these outliers as missing. In this algorithm, I regress each district's crime and police outcomes on a quartic time trend to obtain fitted trend values for each

¹²These groups are determined using the modal population group for a city over the period of 1990-2014.

outcome and then identify outliers that deviate from these fitted values beyond a threshold (see Data Appendix C.2 for more detail).¹³ After cleaning the data, there are 6,966 police districts and 93,081 observations in the violent crime sample and 6,964 districts and 93,296 observations in the property crime sample.

The research design hinges on the grant acceptance and application variables in the model. The variable $ApplyHiring_{it}$ is an indicator for whether a police department applied for a hiring grant in year t , $t - 1$, or $t - 2$, allowing this dummy variable to be set to 1 for the duration of the 3 year grant period for which the agency applied for funding. Similarly, the $AcceptHiring_{it}$ variable is also a dummy variable set to equal 1 if the agency received a grant and is within the 3 year funded grant period. In this structure, a department that applied for funding in 2009 and was rejected would have the indicator $ApplyHiring_{it}$ set to 1 during the 3 year period of 2009, 2010, and 2011 when the department would have received funding if accepted, and the $AcceptHiring_{it}$ variable set to 0 during this same period. Likewise, a department with an accepted grant

¹³The threshold is the maximum of a 50 percent deviation from the fitted value or the 99th percentile of the distribution of deviations from fitted values for cities within a population group. Outliers are identified using a larger sample window of 1990-2014 and are identified separately for violent and property crimes and arrests, allowing these groups of outcomes to have different observations in the model results. The regression algorithm is based on a procedure that is used by [34] to flag outliers for manual visual inspection; in the larger sample in this paper, I use the algorithm to replace observations as missing. The regression algorithm I use to identify outliers is effective at retaining information for small districts that appear to be consistently and accurately reporting crime data. To the best of my knowledge, this paper uses a lower population threshold than other papers in this space, allowing for a larger and more representative national sample. The 1,000 resident population restriction excludes 13% of the cleaned data sample.

in 2009 would have both indicators set to 1 for the 3 year grant period of 2009-2011.¹⁴ The actual grant duration variation that corresponds to the construction of the *ApplyHiring_{it}* and *AcceptHiring_{it}* variables is shown in the lower right graph in Figure 3.1.

Police departments may apply and/or be accepted for multiple grants during the sample period, and over time the same district can change states, alternating between having a rejected grant, an accepted grant, or no application. The total impact of receiving a grant is the sum of *AcceptHiring_{it}* + *ApplyHiring_{it}* while the impact of a grant rejection is *ApplyHiring_{it}*. Because the decision to apply for a grant is non-random and may be influenced by the needs and goals of a police department, *ApplyHiring_{it}* serves as a critical conditioning control in this model, while the conditional *AcceptHiring_{it}* variable is the excluded instrument. The analysis in this paper is focused on the 2000-2014 time period because prior to 2000 nearly all hiring grant applications were accepted (Figure 3.1, lower right graph).

Data on all of the rejected and accepted COPS grants was obtained through a Freedom of Information Act (FOIA) request to the COPS DOJ Office. Through this disclosure, COPS shared information on the grant program,

¹⁴In the UCR data, police officer employment is reported as a snapshot on October 31st. To match this timing feature, grants with start dates for January-October are Indexed to begin in the current calendar year and grants with start dates in November and December are Indexed to begin in the following year (see Data Appendix C.2). I choose to match police and crime variables within the same year because October employment reflects hiring from earlier months in the year, and I observe a strong first stage using this construction. This timing assumption differs from [34] which lags police variables by a year.

project start date, and project end date for both rejected and accepted applications. For accepted grants, the COPS disclosure also included the amount of money awarded and the number of officers eligible for funding through the grant, although this information was often incomplete.¹⁵

The model also includes municipal police district or police department fixed effects, ϕ_i . By including police department fixed effects, the model does not simply compare crime outcomes for districts that received hiring grants to districts that applied for but did not receive a grant. Instead, as in most papers examining the impact of police, the model uses *within* district variation to control for unobserved factors that vary across districts but do not change over time. Using within district variation helps to control for differences in policing culture and objectives, differences in local governments, and differences in crime patterns across cities that are constant over time. The police department fixed effects are critical to the model design, as they also control for baseline differences in police organizations that may affect a department’s likelihood of receiving a grant.¹⁶

¹⁵In their evaluation of the COPS hiring grant program, [34] used variation in the size of the monetary award as an instrument for police force expansions rather than simply whether a police department received a grant. In this paper, it is not possible to use the size of monetary awards as an instrument because comparable information is not available for departments that applied for grants and were rejected. Using variation on the intensive margin could bias the results if the funding amounts of rejected applicants (not observed) systematically differs from the funding amounts requested and received by accepted grantees.

¹⁶Given the inclusion of police department fixed effects, it is not possible to directly include lags of crime rates in the model as this would introduce mechanical correlation between a lagged crime term and the error term.

In this structure, the impact of policing is measured as the difference in crime outcomes in years when a district is receiving federal COPS funding versus years when a district did not apply for funding, while also controlling for changes in policing that occur in years when a district applies for funding and is rejected. Because there is substantial variation in appropriations for hiring grants across years, the number of possible grants that can be funded in a given year is a key driver of the probability of hiring grant acceptance *within* a district over time (Figure 3.1, upper right graph). It is worth emphasizing that the variation in grant receipt *across* districts is more likely to be affected by differences in application effort and other differing attributes across police departments, and that these factors are controlled for by including police department fixed effects, ϕ_i . *The identifying assumption of the model is that conditional on a police department's decision to apply for a hiring grant in a given year, the timing of the acceptance of that department's grant proposal is not a function of changes in crime rates in that police district.*

In addition to satisfying this identifying assumption, the instrument must also satisfy the exclusion restriction. The exclusion restriction requires that the conditional acceptance of a hiring grant may only reduce crime through increasing the size of the police force in a department. In Section 2.4.2 below, I discuss the results from the first stage, which show that the number of police actually hired as a result of a grant is lower than the number of police that are designated to be hired in COPS hiring grant applications. This imperfect pass-through means that some grant funding may serve other purposes

in a police department that cannot be observed or addressed by the researcher. To address this concern, I include controls for COPS grant acceptances and applications for police functions other than hiring, *AcceptOtherGrants_{it}* and *ApplyOtherGrants_{it}*, as hiring grant acceptances may be correlated with both the interest in and receipt of grants for other policing functions that could also influence crime outcomes. Data on grants for non-hiring functions (such as technology, school grants, and community partnership grants) provides a proxy measure of police department interest in investing in these alternative policing functions.¹⁷ After including these controls, the analysis that follows assumes that the transfer of grant funds influence crime through police hiring rather than other means.

2.4 Results

2.4.1 Summary Statistics

The average police district in the sample has 24,278 residents, a police rate of 23.5 officers, a violent crime rate of 34.8, and a property crime rate of 318.1 (per 10,000 residents). The demographics in the sample are comparable to those in the U.S. over this period, with districts that are on average 84% White, 10% Black, and 11% Hispanic, while the average unemployment rate is 7% and the average annual earnings for employed workers was \$35,991

¹⁷These variables are coded to reflect the varying project duration of COPS grants across programs (see Data Appendix C.2).

(Table 2.1).

Within the sample period, 9% of district-year observations are funded within a 3 year COPS hiring grant cycle, while 26% of observations are within a comparable application hiring grant cycle, corresponding to an effective acceptance rate of 35% (Table 2.1). The average hiring award is \$825,585 or \$223,788 per 10,000 residents, designating funding for an average of 5.8 police officers or 2 police officers per 10,000 residents. While only 25% of districts were ever accepted for a hiring grant, 65% of districts applied for at least one hiring grant during the sample period. On average, police departments applied for multiple hiring grants, resulting in 17% of the sample that was both accepted and rejected for a hiring grant within the sample period (Table 2.2).

2.4.2 Primary Results

Table 2.3 shows the primary model results for violent and property crime rates using the preferred specification. The direct OLS estimation of the impact of police on violent and property crime yields small but significant positive coefficients. As discussed above, OLS estimates may be upwardly biased due to simultaneity.

The first stage estimates show that the conditional impact of receiving hiring grant funding results in an increase of 0.65 officers per 10,000 residents, considerably less than the eligible police officer increase designated

by the average grant, which would increase the police rate by 2 officers per 10,000 residents (Table 2.2). These results imply that for every officer funded through a COPS grant, 0.33 officers are actually hired. My first stage estimates are lower than those in [34], who find that for every officer funded through a grant, only 0.7 officers were actually hired. At the same time, the application controls in my model show that police departments that apply for grants and are not accepted actually cut their police forces, and this provides meaningful context to understanding the first stage impact of grant awards.

Though these results are not stated in dollar terms, the gap in eligible hires and actual hires is consistent with the literature on the "flypaper effects" of intergovernmental grants which predicts substantive but imperfect pass-through of targeted grant funding [e.g. 54]. Controlling for grant applications increases the observed conditional grant receipt or implied "flypaper" effects in this paper, because districts that applied and did not receive grants reduced their police staffing levels.

The imperfect pass-through of funding may mean that hires were more costly than projected in the grant applications or that local districts did not comply with the fund matching requirement of grants. Alternatively, these grant funds may have been utilized for other policing purposes that could also impact crime and may be a concern for the exclusion restriction. As discussed above, I have included application and acceptance variables for other COPS grant types as a proxy for a district's interest and investment in non-

hiring policing functions, as these non-hiring functions may be both correlated with hiring grants and affect crime outcomes. Moreover, it is unlikely that departments are reducing crime through diverting grant funds to increased wages for existing officers, because wages and hiring are budget trade-offs and reductions in crime are strongly associated with increasing hiring in the IV estimates.

The coefficient on *ApplyHiring_{it}* is -0.13 to -0.14 and is significant at the 5% level, showing that police departments that applied for and did not receive grants were anticipating staffing reductions. Though this negative coefficient means that the total impact of a grant is less than the conditional acceptance effect, it implies that police departments applying for hiring grants did in fact need these grants to maintain their staffing levels. In contrast, both the *ApplyOtherGrants_{it}* and *AcceptOtherGrants_{it}* coefficients are positive and significant, suggesting that districts that applied for other grant types were planning to increase funding for staffing, whether or not an other grant was received. The instrument *AcceptHiring_{it}* has a high level of predictive power, with associated F-Statistics of 74.6 to 77.5.

In the reduced form models, the impact of receiving a hiring grant on crime rates, conditional on applying, is negative and highly significant. The impact of applying for a hiring grant is positive for both violent and property crime and highly significant for violent crime, showing that police departments that decide to apply for hiring grants subsequently experience

increases in crime rates.

The two stage least squares or instrumental variables estimates in columns (4) and (8) show that police have a large impact on reducing crime. The implied elasticities evaluated at the mean are -1.28 for violent crime rates and -0.73 for property crime rates, and are moderately larger than the -0.3 to -1 range in the prior literature on policing.¹⁸ In this paper, controlling for the elective decision to apply for hiring grants increases the measured impact of policing, because applications are associated with future crime increases. In other words, omitting the application controls would lead to an underestimate of the impact of policing on crime.

In Table B.1, I estimate the impact of different components of the model by successively including additional controls. In this table, all models include controls for hiring grant application and other grant application and acceptance. In columns (1) and (5), the models control only for time fixed effects, and all estimation approaches show a strong positive relationship between police and crime. Adding demographic covariates in columns (2) and (6) does have an impact on the magnitude of these coefficients, but they remain large and positive. When police department fixed effects are added in columns (3) and (7), the sign flips negative for the reduced form and IV regressions, highlighting the importance of using within police district variation

¹⁸Because there are multiple treatments within districts over time, I use calculate elasticities using crime and police means for the entire sample period rather than a "pre-treatment" mean.

rather than cross-sectional variation. In the preferred specification, columns (4) and (8), I add year by city size fixed effects in place of year fixed effects to flexibly control for changing crime trends in rural and urban municipalities. These additional fixed effects modestly reduce the magnitude of the IV estimates.

Using the preferred model, I next estimate the effects of policing on each of the 7 Index I crimes in the UCR data in Table 2.4. These results show that the police impact on the violent crime rate is driven by effects on robbery and aggravated assault. There is no significant impact of police on murder or rape, and these weak results may be due to the fact that these crimes have a low incidence in the sample, making it difficult to identify a change. Within property crime, policing has a large impact on larceny and vehicle theft, but the coefficient on policing in the burglary regression is actually positive with an elasticity of 0.79.¹⁹ Robbery and vehicle theft have the largest elasticities among the crimes with significant effects, with elasticities of -2.1 and -4.1 respectively. The crime effects are most significant for crimes of acquisition: robbery, larceny, and vehicle theft.

¹⁹The positive burglary effect may be an artifact of noise in this variable, as can be seen in the treatment effect over time graph in Table B.1.

2.4.3 Model Validity and Treatment Effects Over Time

As discussed above, the identifying assumption of the model is that conditional on a police department's decision to apply for a hiring grant in a given year, the timing of hiring grant acceptance is not a function of changes in crime rates within a police district. To formally test the identifying assumption, I estimate the treatment effect of hiring grant acceptances and applications interacted with year indicator variables for the three years before and after treatment.²⁰

Figure 2.2 displays the results of these tests, where each row shows the acceptance and application coefficients over time from the same regression on the total sample. The lags for hiring grant acceptances are flat for both violent and property crimes, though in period -3, the lag for violent crime rates is positive and significant. For the application effects, the pre-treatment indicators are also primarily flat for both violent and property crime. There is no observed spike in total violent or property crime in the year prior to treatment for either acceptances or applications, indicating that police departments

²⁰Each treatment is centered within a 6 year time window with the year prior to treatment (year -1) omitted, restricting the sample to hiring applications and acceptances that were submitted between 2000-2012. The years 2000 and 2001 use lagged data that is outside the sample period from 1998 and 1999, that was cleaned and prepared in the same way as data in the sample period. Because districts can have multiple acceptances or applications over time, the data is duplicated and stacked for each year between 2000 and 2012. To control for multiple treatments, I define indicators for hiring and other grant acceptances and applications within each duplicated data set that are defined in all years other than the centered treatment year, when they are set to equal zero. This procedure creates a "pseudo-panel" of duplicated data that functions as if each district had a single treatment.

are not applying for or receiving hiring grants because of transient shocks to crime rates. Despite the inclusion of crime information in COPS hiring grant applications, these tests confirm that prior crime rates were a minor factor in determining acceptances *within* police districts.²¹

Instead of observing pre-treatment deviations in crime conditions for grant applicants, a different picture emerges in these graphs. The acceptance coefficients are negative in the post-treatment period for both violent and property crime and tend to get larger over the grant period. In contrast, the application coefficients are positive and are highly significant for violent crime in the post-treatment period. Consistent with the general model results, this test shows that, on average, police departments that decide to apply for hiring grants subsequently experience increases in crime rates. This pattern suggests that the years that departments choose to apply for hiring funds are "negatively selected." Because of this relationship, failing to control for their application decision would lead to an underestimate of the impact of policing.

In contrast to the crime rate outcomes, it does appear that hiring grant acceptances were more likely to be distributed to districts facing police

²¹Figure B.1 shows the effects for specific crimes. Pre-trends in the timing of conditional acceptance rates (top row of each pane), are insignificant for all sub-crimes except for robbery and burglary, where they are positive. While these effects do not translate to overall crime rate trends, an upward bias in the pre-period could increase the size of the crime elasticities that I estimate. To address these concerns, I include a number of robustness checks in Section 3.4.4 that control for additional aspects of selection in the timing of conditional acceptances, including application cohort time trends, and time trends that vary across cities that were only accepted, never accepted, were both accepted and rejected, or never applied.

staffing shortfalls, as the acceptance coefficients are highly significant and negative in the pre-treatment years when police employment rates are considered as an outcome. This pattern reflects the COPS application consideration of "fiscal need" in grant allocation decisions. Notably, while flat pre-treatment crime trends are necessary for the identification assumptions of the model, a dip in pre-treatment policing levels in the first stage does not violate the assumptions of the estimation approach. Consistent with evidence from the first stage of the model, the bottom right graph of Figure 2.2 shows that police departments applying for grants were also anticipating moderate reductions in staffing.

Though prior research on policing discusses the issue of the joint determination of crime and policing in a given time period, this research has not been able to empirically investigate whether police departments may respond to their outlook of future crime. If police hiring is not only a function of current or prior crime rates but is also a function of expectations of crime rates in later periods, there are additional measurement obstacles to understanding the causal impact of police beyond simple simultaneity bias. The relationship between police application decisions and subsequent crime increases could in part be due to foreseeable police staffing shortfalls that contribute to crime increases. At the same time, the hiring application coefficients in this paper provide suggestive evidence that police departments may have an understanding of their communities that allows them to make dynamic hiring decisions based on expectations of changes in crime.

The relationship between police department grant application decisions and subsequent crime rates is also consistent with theories of fiscal federalism [e.g. 87], which investigate optimal assignment of taxation and spending across different tiers of government. In this context, police departments have local and specific knowledge about the crime conditions in their communities as well as the resource needs of their organizations. When departments seek funding and are not assisted by the federal government, they appear to operate at a lower level of efficiency.

2.4.4 Robustness Tests

Next, I perform several robustness tests of the model. First, I conduct placebo tests that randomly vary the timing of hiring grant acceptance among police departments that apply for hiring grants, to show that the observed effects are due to actual hiring grant receipt and not due to chance. Figure 2.3 maps the distribution of 1,000 replications of these randomized timing placebo tests and shows that the model estimates are well below the 95% confidence interval of the replication distribution.²²

Table 2.6 shows the results of additional robustness tests using different analysis samples and methodologies.²³ Specification 1 extends the

²²I assign placebo acceptances in such a way that the placebo acceptance rate matches the observed acceptance rate over the total sample period. Results are similar for placebo tests that randomize both the application timing and acceptance timing for hiring grants.

²³I have also explored the heterogeneity of the police impacts by interacting the key police rate coefficient with characteristics of police districts but found few interesting patterns

sample period to 1990-2014 to include years when nearly all hiring grants were accepted and finds smaller impacts of hiring grants, with no significant impacts for larceny and aggravated assault. Because it is not possible to effectively control for the application effect between 1990-2000, this larger sample period may underestimate the impact of police by not incorporating the fact that departments face future crime increases when they apply for grants.

In Specification 2, I expand the sample to include small police departments with fewer than 1,000 residents, a sample that may have a higher rate of reporting errors. The results in this sample are directionally consistent with those from the base analysis sample, though the coefficients are smaller in the noisier sample. When the model observations are weighted by population in Specification 3, the coefficients increase in magnitude, showing that the crime reduction effects are driven by districts with larger populations. In these tests, the elasticity for violent crime is -2.7 and for property crime is -2, over 2 times the elasticities in the base model.

Because criminal justice policies and policing tactics may have common elements within states, I include state by year fixed effects in Specification 4, replacing the population group by year fixed effects in the base model. This specification allows crime time trends to differ across states, rather than

beyond stronger impacts in more populous cities (results available on request). In these tests, crime reduction effects appear to be larger in districts that have a greater proportion of non-white residents, however, this result may be partly due to the fact that diverse districts are also likely to be more populous and effects are stronger in larger cities.

across different city population groups (base model). The estimates using this approach are quite similar to the base specification.

In Specifications 5 and 6, I consider alternatives to the data cleaning approach that is used in the analysis. In Specification 5, I use an alternative method to identify outliers. Here, I define the ratio of crime outcomes (and police employment) to the within city mean for a variable and then remove observations with ratios that are below the 1st percentile and above the 99th percentile within each population group. This cleaning procedure drops more observations than my preferred algorithm, yielding a sample of 65,548 for violent crime and 74,663 for property crime. The results using this procedure are consistent with the primary results, and are stronger for larceny and vehicle theft, while the positive coefficient on burglary is also larger and more significant.²⁴ In Specification 6, I expand the sample to include the raw data that does not exclude the outliers that were identified when the data was cleaned.²⁵ These results show the importance of cleaning the data as none of the crime outcomes show notable or significant impacts of police on crime reduction in this specification.

In the second part of Table 2.6, I formulate a number of robust-

²⁴Like the main cleaning algorithm, this procedure identifies outliers for groups of outcomes, violent crime, property crime and police employment, and then sets sub categories equal to zero when there is an outlier in the larger group. This procedure also uses data from 1990-2014 to identify outliers.

²⁵This raw data does include basic data cleaning but was not subject to the algorithm to identify outliers (see Data Appendix C.2).

ness checks related to the grant variation I use in the estimation strategy. First, I directly consider the impact of controlling for application by estimating the model without these controls in Specification 7. These estimates are less negative and less significant than the preferred model, showing that failing to condition on application decisions leads to an underestimate of the crime reducing impact of police hiring.

I construct an illustrative test in Specification 8 that uses award dollars per capita as an instrument for police employment rates, a strategy similar in spirit to the methodology in [34]. To do this, I exclude municipalities with missing award information for accepted hiring grants or other COPS grants at any point in the sample period. Unfortunately, I do not have information on the award amount requested for grants that were rejected, so these tests do not contain controls for grant applications. Like Specification 7, these estimates are weaker and less significant than the preferred model. The elasticities in this model are -0.62 for violent crime and -0.46 for property crime, which are close to those reported in [34] of -0.99 and -0.26, respectively.

Restricting the sample to police agencies that applied for a hiring grant in Specification 9, or to agencies that were accepted for a hiring grant in Specification 10, does little to change the magnitude or general pattern of results. In Specification 11, I further restrict the sample to include only those agencies that were both rejected for a hiring grant and received a hiring grant in the sample period. Despite shrinking the sample by over 60%, these results

are roughly consistent with the main model estimates, though many of the coefficients are not significant.

A small fraction of grant recipients withdrew their applications and returned their funding after their project was accepted by the COPS Office (13% of accepted grant years). In the main model, I consider withdrawn acceptances within the accepted category, which renders the grant coefficients in the first stage as "intent to treat" estimates. In Specification 12, I consider withdrawn hiring and other grant applications as a separate category and exclude them from accepted grants. The results are highly consistent with the preferred specification, showing that the categorization of these grants is not critical to the analysis.

In Specification 13, I allow the crime time trends in police districts to vary by their hiring grant acceptance profile. I define four mutually exclusive groups of police departments, (1) departments for which all applications were accepted, (2) departments for which all applications were rejected, (3) departments that were both accepted and rejected on COPS grant applications, and (4) departments that did not apply for grants. I allow the crime trends to differ across these acceptance groups by inserting acceptance group by year fixed effects in place of the population group by year fixed effects in the base model. The effects in this specification are comparable to the base model, with effect sizes that are somewhat larger.

Lastly, in Specification 14, I allow the year fixed effects to vary

by grant application cohort, interacting the year effects with whether a police department submitted a new application for a grant in that year.²⁶ As in Specification 4, this estimation replaces the population group by year fixed effects with an alternate set of year fixed effects. This model allows time trends to vary by a department's time-varying application decisions. The results from this robustness check are consistent with the base model, though they are larger and more significant.

2.4.5 Mechanisms: Crime Clearances

While prior work has consistently found that police reduce crime, there is less consensus about the mechanisms behind this effect.²⁷ Broadly, when police reduce crime rates, they may do so by taking more individuals into custody and increasing total levels of incapacitation, or by reducing crime through deterrence. The distinction between incapacitation and deterrence is important because incapacitation entails additional costs beyond the base cost of hiring police, including the direct costs of court proceedings and incarceration.

²⁶I define application cohorts using the first year that a district applied for a grant, going back to the beginning of program data in 1994. Due to small numbers of "first applications" in certain years, I group 2004-2008 cohorts and the 2010-2014 cohorts together.

²⁷There is a growing literature that evaluates particular police tactics and finds mixed results. For example, there is evaluation evidence that "hot spots" policing, a tactic that focuses policing resources on high crime areas within a city, is an effective approach to reducing crime [e.g. 9, 101], while other papers have found that "broken windows" policing, which emphasizes heavy enforcement for low level crimes, has weak crime reducing effects [12]. Likewise, advances in computing and IT in police departments may reduce crime when they are coupled with complementary management strategies, including personnel training and specialization [43].

tion and indirect "collateral" costs to incarcerated individuals.²⁸

It may also be the case that while crime falls, arrests stay flat or decrease at a slower pace than crime, meaning that while total incapacitation levels do not increase, the proportion of crimes that result in apprehension increases. A recent book chapter by [89] finds that arrests for Index I crimes do not increase as a result of COPS hiring grants, and because of this, she attributes the decline in crime caused by police hiring to be due to increased deterrence from a greater policing presence. A growing literature in economics and criminology finds that the certainty of punishment may be more important to deterring potential offenders than punishment severity [53, 68, 73, 83]. If this is true, an increase in crime solving rates may be tied to an increase in deterrence and reductions in crime.

Criminologists often examine clearance rates, or the number of crimes solved or "cleared" (by arrest or other means) divided by the total number of reported crimes, as a metric of police performance in crime solving.²⁹ When modeling clearances as an outcome, it is important to contextualize changes in clearances, as clearance rates are not direct measures of police effectiveness [19]. For example, a district that has a constant or declining

²⁸There is extensive literature on the collateral costs of incarceration in the United States. These costs may include increased rates of reoffending, lost income, higher rates of unemployment, and adverse impacts on families and children [e.g. 59, 80, 90, 92].

²⁹Cleared or solved crimes are correlated with arrests and changes in incapacitation and are available for Index I crimes in the UCR data. Cleared crime counts are not directly comparable to reported crimes because they may include crimes solved that were reported in earlier years. A crime may also be cleared without an arrest (e.g. a suspect is deceased).

clearance rate and a decline in crime may be experiencing an increase in policing productivity. While examining clearance rates is the standard approach to estimate the impact of additional police on crime solving performance, I do not use clearance rates as an outcome in this paper because clearance rates are undefined when the number of reported crimes in a category is zero.

Instead, I estimate the impact of police rates on total numbers of cleared crimes adjusted per 10,000 residents in Table 2.5. I then conduct a bootstrap sample test to measure whether the declines in reported crime and cleared crime are proportional by testing if their estimated elasticities are equal.³⁰ For the categories with significant reductions in total reported crime rates, the measured elasticities for cleared crimes are nearly all less negative than their corresponding crime rates, and are significantly less negative for robbery, aggravated assault, larceny, total violent crime and total property crime. This pattern suggests that while the absolute number of crimes cleared does not increase when more police are hired, the number of crimes cleared increases relative to the number of crimes reported. Because changes in total cleared crimes are a measure of changes in incapacitation, total incapacitation levels are likely decreasing while crime solving rates are also going up. These tests suggest that larger police forces are better equipped to solve crime and may increase deterrence by raising the probability of apprehension and

³⁰To construct this test, I draw 500 bootstrap samples from the data and derive the distribution of differences in the elasticities from estimating the models on each sample. I then determine the probability of the original observed elasticity difference relative to the bootstrap distribution of differences under the null hypothesis.

punishment for criminal behavior.

2.4.6 Shifts in Policing Focus

To augment the cleared crime analysis, I include illustrative outcomes of arrests per 10,000 residents in Table 2.7 in order to gain an understanding of how policing activity may change when more police are hired.³¹ Aside from a slight decline in vehicle theft, there are no significant changes in arrests for Index I crimes adjusted for population, consistent with the results for cleared crimes in Table 2.5 and the findings in [89]. However, there are significant changes in arrest outcomes for categories outside of the 7 Index I crimes. A greater police presence is associated with increases in arrests for marijuana sale, while it is also associated with decreases in arrests for the sale and possession of narcotics. The impact of police expansions on DUI arrests is also positive and significant, with an elasticity of 1.2.

Unfortunately, arrest outcomes for drug crime and intoxicated driving cannot easily be benchmarked to crime rates because reported crime rates are not available for these arrest categories. Consequently, it is not possible

³¹Arrest counts are available for a larger set of crimes in the FBI UCR data, including several crimes that are not reported by victims, including drug crimes and driving under the influence (DUI). I chose to add the categories of simple assault, drug, and DUI arrests because they are high frequency arrest categories that are more likely to be consistently and accurately reported. While an arrest may be a means of clearing a crime, arrests are counted in units of people apprehended and are not directly comparable to crime units because more than one person may be arrested for a crime. Arrest counts are included at the person-incident level, with the most serious arrest charge recorded for each person-incident.

to differentiate between a decrease in narcotics arrests that corresponds to a decline in use of narcotics due to increased policing or a decrease in narcotics arrests that is associated with a reduced emphasis on narcotics enforcement. For arrest outcomes that increase, it is less likely that these arrest spikes are linked to a corresponding increases in crime, given that the timing of hiring grant acceptances is not related to pre-treatment levels of reported crimes that are observable. Instead, increases in marijuana and DUI arrests are more likely to occur as a result of increased policing capacity and a shift in policing focus. The shifts in arrests provide evidence that police departments change tactics and enforcement patterns when police forces expand.

I also explored how police hiring affects arrests for different demographic subgroups in Table B.2 and Table B.3. The decline in vehicle theft rates is most significant for Blacks and young adults aged 15-24. Changes in drug arrests (both the increases in marijuana arrests and decreases in narcotics arrests) are more pronounced for Blacks, males, and young adults aged 15-24. The more elastic arrest responses for young Black men may indicate that newly hired police are focused on policing this demographic group, possibly through increasing patrols in Black neighborhoods. At the same time, reductions in larcenies and simple assault arrests are strongest for children younger than 15 and increases in DUI arrests are more significant among Whites and males.

2.5 Conclusion

This paper uses a novel identification strategy to measure the causal impact of policing on crime. Using a research design that exploits variation in federal COPS hiring grants applications and acceptances, I find that an increase in police presence significantly reduces crime, with elasticities of -1.28 for violent crime and -0.73 for property crime.

Given the large crime reducing effects of police hiring on serious crime, investments in police force expansions are likely to be cost-effective. Leveraging the results in this paper, a "back-of-the-envelope" calculation can be used to illustrate the welfare impact of police hiring in the context of Index I crimes. If the average district in the sample increased its police employment rate by 1 officer per 10,000 residents, or 2.4 total additional police officers, the primary results imply that there would be an annual reduction of 1.5 robberies, 3 aggravated assaults, 20.9 larcenies, and 8.5 vehicle thefts, while increasing the number of burglaries by 5.3. The cost of this police hiring increase can be determined directly from the summary statistics and first stage estimates. The average hiring grant award was \$258,234 over 3 years (per 10,000 residents) and resulted in an increase in the police rate of 0.65. This implies that the annual cost of raising the police rate by 1 officer per 10,000 residents is \$132,428, or \$54,546 per individual officer-year. Applying upper and lower bound estimates of the literature on the social cost of crime, I find that the average welfare

benefit of this hiring increase ranges from -\$16,007 to \$556,306 per district.³² Though this range is wide, this exercise confirms that funding for police hiring likely passes a cost-benefit test when considering changes in Index I crimes.

The clearance and arrest analysis also have welfare implications for policies that expand police presence. The cleared crimes analysis shows that police hiring raises the certainty of apprehension without increasing total incapacitation, providing evidence that the declines in Index I crime rates resulting from police hiring are caused by deterrence. Because total incapacitation rates for these crimes decrease when police forces expand, any additional costs related to incarcerating offenders are not likely to outweigh the welfare benefits of reducing Index I crime rates.

Outside of Index I crimes, investments in police hiring do have significant impacts on other arrest types, decreasing arrests for narcotics offenses and increasing arrests for marijuana offenses and DUI crimes. These arrest

³²All monetary values are computed in 2015 dollars in this exercise, using the CPI Inflation Index from the Bureau of Labor Statistics. The lower bound estimates are derived from [78] that calculates social crime costs by estimating the different components of tangible and intangible victim costs. The upper bound estimates are derived from [18] that estimates social crime costs from survey responses about willingness-to-pay to avoid crime. Because larceny and vehicle theft are not included in [18], I calculate the upper bound for these crimes as the sum of the lower bound and the value of property lost (or "property transfer") for the average crime in each category using national data from the FBI UCR in 2015, following the approach in [80]. The cost-benefit calculations shown here do not consider additional deadweight loss resulting from government transfers or multipliers of employment. I have also excluded local police department salary matching costs because the first stage estimates suggest that far fewer police officers are actually hired by the grants than are designated by the grants. This shortfall suggests that police departments are not sufficiently matching hiring grant funds on average.

changes are difficult to translate to total welfare because they cannot be contextualized to underlying crime rates for these offenses. If one presumes that declines in narcotics arrests are indicative of reductions in narcotics crime, the welfare gains for this crime group are likely positive, decreasing total narcotics crime and incapacitation for these offenses. Alternatively, the welfare impact of increases in arrests for marijuana offenses and DUIs are more uncertain. These welfare effects depend on the underlying changes in the incidence of these offenses, the social cost of these offenses, and the costs and benefits of additional incapacitation of offenders that are apprehended. Further, because increases in arrests for different offense types are concentrated among particular demographic groups that may be geographically segregated (e.g. young Black males for marijuana sales and White males for DUI offenses) there may be additional community welfare consequences of police hiring.

Beyond providing evidence that police force expansions reduce crime and are likely cost-effective, this paper offers insight into the nature of simultaneity bias in measuring impacts of police. By mapping how crime rates change in districts that apply for hiring grants, I observe that the decision to apply for federal hiring funds is not motivated by a spike in crime, but is instead correlated with subsequent increases in crime rates. This pattern reveals that police departments may not only be responsive to current crime conditions, but may also react to expected changes in crime. This paper is the first to document suggestive evidence of that police decisions are forward-looking, and this behavior highlights additional challenges in measuring the causal im-

pact of police. Moreover, this empirical relationship is indicative of specific knowledge that police departments may have about local crime conditions and resource needs.

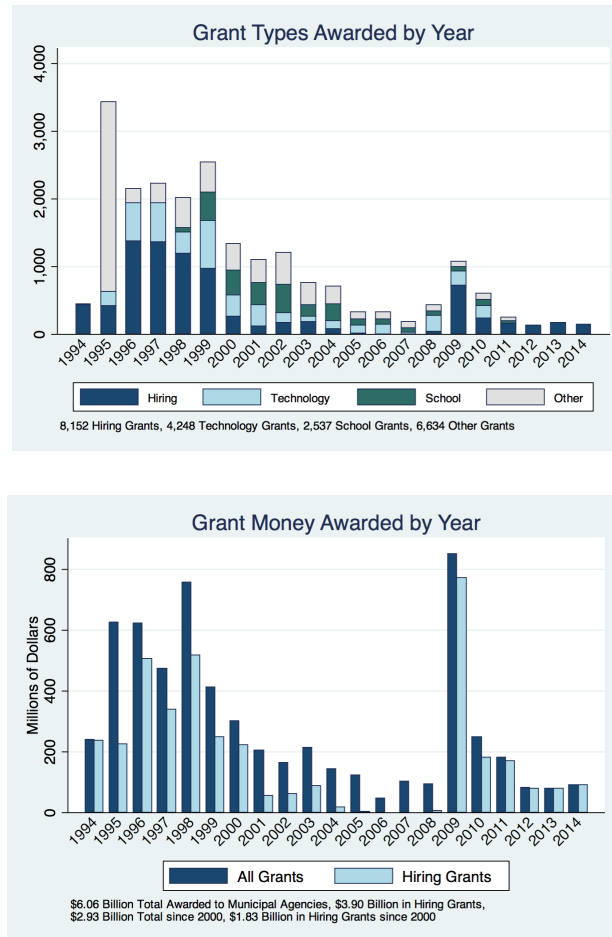
While this paper demonstrates that police hiring has a substantial impact on crime reduction, the results do not provide information on the effect of police tactics on other critical outcomes, including racial bias or use-of-force.³³ Investments in police should not only improve the safety of communities but should also establish trust between communities and law enforcement. More research is needed to understand best practices in policing as well as the welfare implications of different policing approaches.

³³A large body of literature in economics, sociology, and criminology finds evidence that minority individuals are more likely to interact with the criminal justice system and face higher probabilities of arrest. In economics, studies have found evidence of statistical discrimination [6, 5] and taste-based discrimination [104] in police interactions.

2.6 Tables and Figures

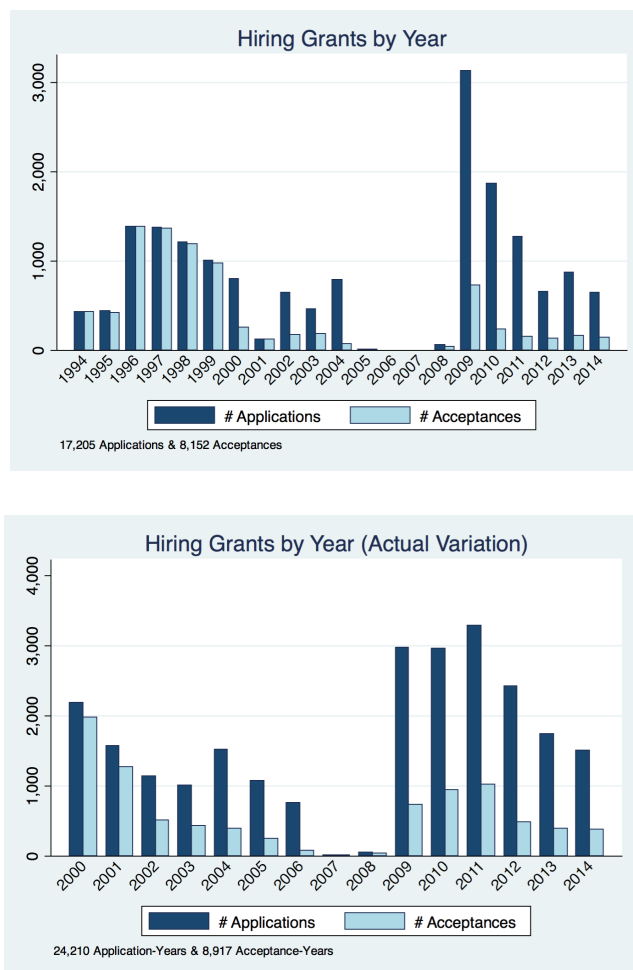
Figure 2.1: COPS Grant Program over Time

Figure 2.1.A: COPS Grant Program, Types of Grants



In these graphs hiring grants are attributed to the years in which they were applied for or awarded. The COPS grants tabulated here correspond to the sample used in analysis rather than the universe of grants applied for and awarded. In the top graph, grant counts are collapsed within a category, year, and police department, such that an agency that receives two new technology grants in a given year is only counted once. Likewise, in the rare cases that there is more than one new hiring application for the same agency in the same year, these grants are also collapsed, and outcomes are considered acceptances if one or more grants is accepted. The sample consists of municipalities with over 1,000 residents and is also restricted by availability of population size, demographic covariates, and censoring of outcome outliers. Funding sharply decreased for the COPS office between 2000 and 2008, reaching a low of approximately \$210 million in new grants (of all types) distributed in 2007, with no new hiring grants distributed in 2006 or 2007. The data shown above for 1994 to 1999 uses comparable sample descriptions as the actual sample period.

Figure 2.1.B: Hiring Grants, Applications and Acceptances



In the top graph, hiring grants are attributed to the years in which they were applied for or awarded. In the lower graph, the actual variation used in the analysis is shown in terms of grant-years. In this graph, each hiring grant has a project period of 3 years that begins in the first year of the grant. Application periods are also defined to last 3 years. The COPS grants tabulated here correspond to the sample used in analysis rather than the universe of grants applied for and awarded. Grant counts are collapsed within a category, year, and police department, such that an agency that receives two new technology grants in a given year is only counted once. Likewise, in the rare cases that there is more than one new hiring application for the same agency in the same year, these grants are also collapsed, and outcomes are considered acceptances if one or more grants is accepted. The sample consists of municipalities with over 1,000 residents and is also restricted by availability of population size, demographic covariates, and censoring of outcome outliers (see Data Appendix C.2). Funding sharply decreased for the COPS office between 2000 and 2008, reaching a low of approximately \$210 million in new grants (of all types) distributed in 2007, with no new hiring grants distributed in 2006 or 2007. The actual analysis sample is restricted to 2000-2014 though the grant program was established in 1994. The data shown above for 1994 to 1999 uses comparable sample descriptions as the actual sample period.

Table 2.1: Summary Statistics, Covariates and Outcomes

	Mean	S.D.
Number of Law Enforcement Agencies	6,990	
Number of Observations	94,264	
<i>Covariates</i>		
Population	24,278	(137,612)
% Population <14 Years Old	0.20	(0.03)
% Population 15-24 Years Old	0.14	(0.03)
% Population 25-39 Years Old	0.19	(0.03)
% White	0.84	(0.14)
% Black	0.10	(0.13)
% Hispanic	0.11	(0.14)
% Male	0.49	(0.02)
Unemployment Rate	0.07	(0.03)
Annual Pay	35,991	(9,948)
<i>Hiring Variables</i>		
Police Rate	23.53	(12.88)
Accept Hiring Grant	0.09	(0.29)
Apply Hiring Grant	0.26	(0.44)
Accept Other Grant	0.12	(0.33)
Apply Other Grant	0.18	(0.39)
<i>Violent Crime Rates</i>		
Violent Crime	34.80	(39.09)
Murder	0.32	(1.05)
Rape	2.63	(3.82)
Robbery	6.86	(11.56)
Aggravated Assault	24.99	(30.56)
<i>Property Crime Rates</i>		
Property Crime	318.1	(238.2)
Burglary	65.49	(57.25)
Larceny	232.4	(185.3)
Vehicle Theft	20.23	(26.22)

Crime outcomes and police variables are per 10,000 residents in a municipality. Because of the procedure used to identify outliers, violent crimes and property crimes have different numbers of observations used in analysis. Violent crime outcomes include 6,966 agencies and 93,081 observations. Property crime outcomes include 6,964 agencies and 93,296 observations. Covariate and grant variable statistics are calculated using both sets of observations, or on 6,990 agencies and 93,296 observations. Summary statistics refer to observations used in the models, over the period of 2000-2014. The grant variables shown in this table reflect the grant variation used in the analysis with variables constructed to cover the duration of a grant project or the period that a grant project is intended to occur. This duration is 3 years for hiring grants and varies for other grant types.

Table 2.2: Grant Characteristics

	<i>Accept</i>		<i>Apply</i>	
	Mean	S.D.	Mean	S.D.
Number of New Grant Years	6,174		18,190	
<i>Grant Type (Proportion of Column)</i>				
Hiring	0.39	(0.49)	0.62	(0.48)
Technology	0.31	(0.46)	0.16	(0.36)
School	0.33	(0.47)	0.20	(0.40)
Other Type	0.39	(0.49)	0.28	(0.45)
<i>Accepted Hiring Grants</i>				
Hiring Award (\$)	825,585	(3,552,003)		
Hiring Award (\$) per 10,000 Residents	223,788	(227,590)		
Eligible Hires	5.80	(30.29)		
Eligible Hires per 10,000 Residents	2.01	(2.47)		
<i>By Law Enforcement Agency</i>				
	Mean	S.D.		
Number of Law Enforcement Agencies	6,990			
<i>Hiring Grants</i>				
Ever Applied	0.65	(0.48)		
Number of Applications	1.62	(1.68)		
Ever Accepted	0.25	(0.43)		
Number of Acceptances	0.34	(0.71)		
Ever Rejected	0.58	(0.49)		
Number of Rejections	1.28	(1.49)		
Both Accepted and Rejected	0.17	(0.38)		

The number of new grant years is a count of agencies that applied for or were accepted for any new grants in each year, summed over the sample period. Grant type statistics are proportions of acceptances (6,174 new grant-years) and applications (18,190 new grant-years) in each respective column. An agency may receive multiple grants in a year of different types, allowing the grant type columns not to sum to one. In some cases, an agency may apply or be accepted for more than one grant within a type. Across the analyses in this paper, only collapsed types are considered. In other words, if an agency has one or more other grants in a year, that agency has the *AcceptOtherGrants_{it}* indicator variable turned on. Likewise, the number of applications is a count of hiring grants collapsed by year. Total award sums all funds received across grant types and are summed across all observations with any grants awarded. The total award mean is lower than the hiring award mean because it covers a larger set of grant-years and includes observations with only other grants. The number of law enforcement agencies are those with either data for property or violent crime outcomes that are used in the analyses, or 6,990 agencies over the period 2000-2014.

Table 2.3: Impact of Police on Crime

	Violent Crime				Property Crime			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	Police Rate	Reduced Form	IV	OLS	Police Rate	Reduced Form	IV
<i>Key Variables</i>								
Police	0.263*** (0.048)			-1.892* (0.75)	2.173*** (0.405)			-9.904** (3.808)
Accept Hiring		0.648*** (0.073)	-1.227** (0.461)			0.658*** (0.073)	-6.52** (2.331)	
<i>Application Controls</i>								
Apply Hiring	0.746** (0.256)	-0.139** (0.054)	1.234*** (0.297)	0.972*** (0.279)	-0.311 (1.343)	-0.133* (0.054)	2.383 (1.54)	1.064 (1.488)
Accept Other Grant	0.06 (0.509)	0.367*** (0.086)	0.402 (0.516)	1.097+ (0.644)	2.53 (2.588)	0.346*** (0.087)	4.68+ (2.650)	8.106* (3.371)
Apply Other Grant	0.776+ (0.455)	0.225** (0.08)	0.704 (0.457)	1.13* (0.494)	2.241 (2.35)	0.226** (0.081)	1.993 (2.381)	4.234 (2.618)
% Effect		2.75%	-3.53%			2.79%	-2.02%	
Elasticity	0.178			-1.28	0.161			-0.733
F-Test: Accept Hiring		74.57				77.52		
Y Mean	34.8	23.53	34.8	34.8	323.1	23.62	323.1	323.1
Observations	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
Number of Departments	6,966	6,966	6,966	6,966	6,964	6,964	6,964	6,964

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. Each specification corresponds to the preferred covariate set, including demographic covariates, police department fixed effects and year by city size fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants.

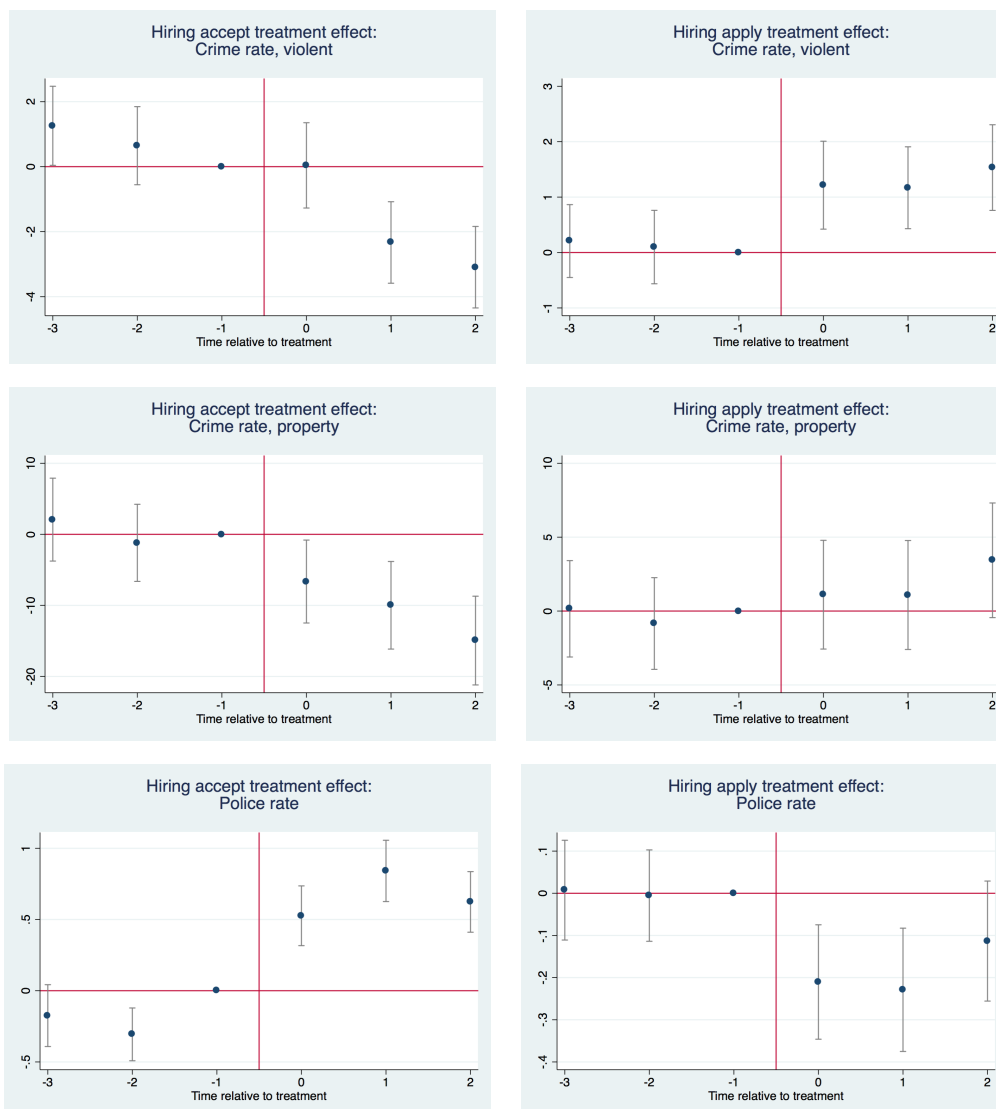
Table 2.4: Impact of Police on Crime, Specific Crimes

	Violent Crime				Property Crime				
	(1) Violent Crime	(2) Murder	(3) Rape	(4) Robbery	(5) Aggravated Assault	(6) Property Crime	(7) Burglary	(8) Larceny	(9) Vehicle Theft
Police	-1.892* (0.75)	-0.007 (0.022)	-0.024 (0.086)	-0.613**** (0.186)	-1.249+ (0.651)	-9.904** (3.808)	2.194* (0.984)	-8.606** (3.107)	-3.493*** (0.674)
Elasticity	-1.28	-0.475	-0.219	-2.102	-1.176	-0.733	0.789	-0.872	-4.065
Y Mean	34.8	0.322	2.626	6.858	24.99	318.1	65.49	232.4	20.23
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
Number of Departments	6,966	6,966	6,966	6,966	6,966	6,964	6,964	6,964	6,964

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

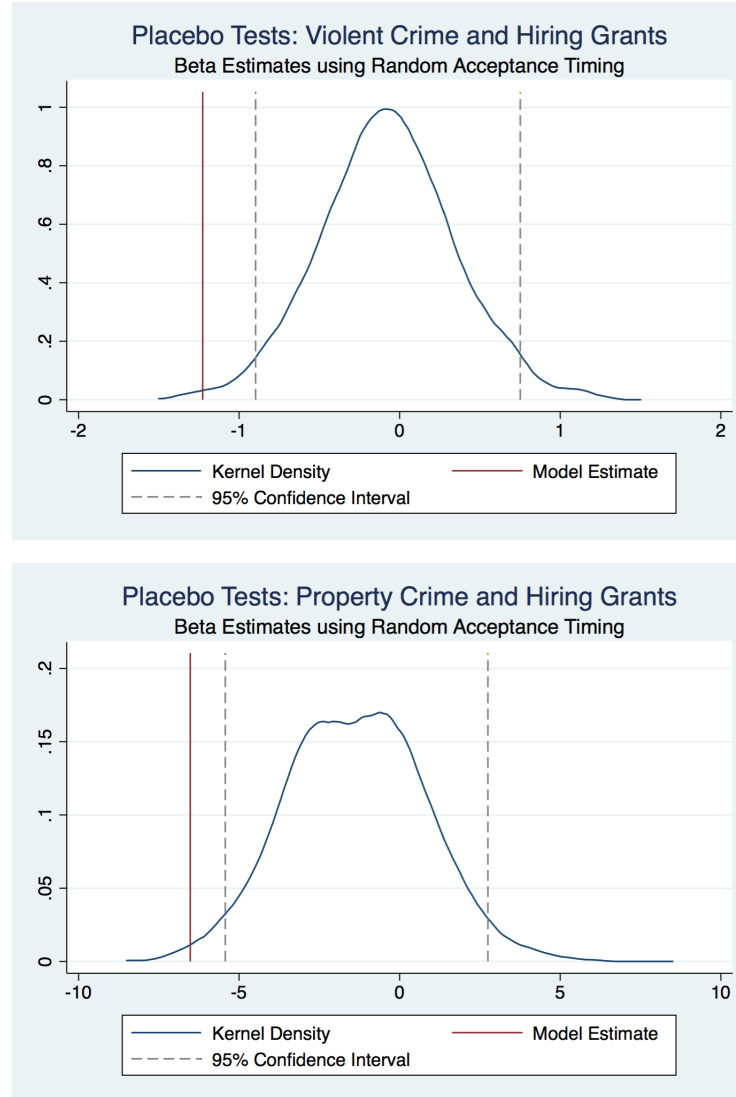
Standard errors in each model are robust and clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. Each specification corresponds to the preferred covariate set, including demographic covariates, police department fixed effects and year by city size fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and acceptances and acceptances of other COPS grants.

Figure 2.2: Timing of Treatment, Application and Acceptances



The range around each estimate represents a 95% confidence interval using robust standard errors clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. The specification shown is the preferred specification of the reduced form model. Coefficients in the acceptance and application graphs in each row are derived from the same regression on the total sample of municipal police districts. Graphs are created by duplicating balanced panels of 6 years corresponding to a centered treatment year for each year between 2000-2012. Specifically, period dummies for hiring application (coefficients shown on right) and hiring acceptance (coefficients shown on left) are included for each period, with the year prior to the first treatment year omitted as a reference year. Treatments are Indexed to the first year of a new hiring grant acceptance or application (year 0 above). Because police agencies may have multiple and overlapping treatment periods, duplication allows the model to control for other coinciding grant treatment effects in years before and after the new treatment.

Figure 2.3: Placebo Timing Tests, Randomized Attribution of Grant Applications and Acceptances (1,000 Replications)



These graphs show the kernel density distributions of 1,000 replications of the model where the timing of grant applications and acceptances is randomized. The randomization procedure is restricted to police agencies that applied for a hiring grants during the sample period of 2000-2014. Hiring acceptances are randomly attributed to agencies that applied for hiring grants in the years that they applied, in such a way that the acceptance rate in the placebo replication is the same as in the observed sample. The confidence interval shown is based on the mean and standard deviation of the distribution of coefficients from the 1,000 replications. In each of the placebo replication models, the preferred fully-specified model is used with robust standard errors clustered at the police department level.

Table 2.5: Crime Clearances as a Mechanism, Reported vs. Cleared Specific Crimes

	Violent Crime				Property Crime				
	(1) Violent Crime	(2) Murder	(3) Rape	(4) Robbery	(5) Aggravated Assault	(6) Property Crime	(7) Burglary	(8) Larceny	(9) Vehicle Theft
Reported Crimes									
Police	-1.892* (0.75)	-0.007 (0.022)	-0.024 (0.086)	-0.613*** (0.186)	-1.249+ (0.651)	-9.904** (3.808)	2.194* (0.984)	-8.606** (3.107)	-3.493*** (0.674)
Elasticity	-1.28	-0.475	-0.219	-2.102	-1.176	-0.733	0.789	-0.872	-4.065
Cleared Crimes									
Police	-0.317 (0.471)	-0.008 (0.017)	-0.042 (0.052)	-0.028 (0.072)	-0.239 (0.433)	1.106 (1.596)	0.863* (0.364)	0.674 (1.325)	-0.430** (0.166)
Elasticity	-0.411	-0.851	-0.942	-0.307	-0.382	0.391	2.042	0.302	-2.460
Difference in Elasticities	-0.869+	0.376	0.723	-1.795*	-0.794+	-1.124+	-1.253	-1.174+	-1.605
P-value (Bootstrap): Elasticities Equal	0.056	0.814	0.411	0.044	0.080	0.052	0.124	0.064	0.104
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
Number of Departments	6,966	6,966	6,966	6,966	6,966	6,964	6,964	6,964	6,964

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. All models shown have the preferred set of covariates, including demographic covariates, application and acceptance controls for other grants, year by city size fixed effects and police department fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. Cleared crimes are defined as the number of crimes that result in arrest or are solved through other means in a given year (also expressed as rates per 10,000 residents in a municipality). The test for the equality between these elasticities was conducted by drawing 500 bootstrap samples (with replacement) and looking at the distribution of differences in elasticities centered at zero. The p-values reported are derived from this bootstrap distribution.

Table 2.6: Robustness Tests and Samples, Specific Crimes (Part 1)

	Violent Crime					Property Crime			
	(1) Violent Crime	(2) Murder	(3) Rape	(4) Robbery	(5) Aggravated Assault	(6) Property Crime	(7) Burglary	(8) Larceny	(9) Vehicle Theft
<u>Base Model</u>									
Police	-1.892*	-0.007	-0.024	-0.613***	-1.249+	-9.904**	2.194*	-8.606**	-3.493***
	(0.75)	(0.022)	(0.086)	(0.186)	(0.651)	(3.808)	(0.984)	(3.107)	(0.674)
Elasticity	-1.28	-0.475	-0.219	-2.102	-1.176	-0.733	0.789	-0.872	-4.065
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
1. Expanded Period: 1990-2014									
Police	-0.873+	0.002	0.048	-0.336**	-0.587	-3.602	1.288*	-3.149	-1.742***
	(0.458)	(0.011)	(0.045)	(0.109)	(0.402)	(2.421)	(0.585)	(1.946)	(0.351)
Elasticity	-0.529	0.126	0.441	-1.029	-0.489	-0.239	0.413	-0.287	-1.717
Observations	148,379	148,379	148,379	148,379	148,379	148,756	148,756	148,756	148,756
2. Adding Districts with Population<1,000									
Police	-0.808	-0.01	0.006	-0.279*	-0.525	-8.395*	0.50	-6.175*	-2.719+
	(0.495)	(0.011)	(0.038)	(0.121)	(0.395)	(3.98)	(0.499)	(2.966)	(1.411)
Elasticity	-0.642	-0.858	0.0611	-1.102	-0.581	-0.682	0.211	-0.676	-3.38
Observations	106,930	106,930	106,930	106,930	106,930	107,151	107,151	107,151	107,151
3. Weighted by Population									
Police	-7.297*	0.032	-0.014	-2.882**	-4.433+	-36.09*	4.80	-27.99*	-12.9*
	(3.078)	(0.131)	(0.158)	(1.077)	(2.391)	(14.71)	(3.505)	(11.38)	(5.196)
Elasticity	-2.742	1.02	-0.101	-3.194	-2.791	-2.01	1.304	-2.287	-6.337
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
4. State by Year Fixed Effects									
Police	-1.879*	-0.007	-0.026	-0.779***	-1.067+	-8.277*	2.678**	-6.574*	-4.381***
	(0.735)	(0.022)	(0.084)	(0.197)	(0.630)	(3.688)	(0.978)	(2.987)	(0.736)
Elasticity	-1.270	-0.526	-0.230	-2.672	-1.005	-0.612	0.963	-0.666	-5.098
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
5. Alternate Outlier Cleaning Procedure									
Police	-1.808*	-0.053	-0.005	-0.750**	-1.001	-14.263**	3.980**	-13.027**	-5.215***
	(0.886)	(0.035)	(0.112)	(0.289)	(0.743)	(4.742)	(1.391)	(4.0)	(1.02)
Elasticity	-1.050	-3.175	-0.0392	-1.999	-0.833	-0.963	1.314	-1.204	-5.387
Observations	65,548	65,548	65,548	65,548	65,548	74,653	74,653	74,653	74,653
6. Raw Data Set, Outliers included									
Police	2.694	0.035	0.04	0.859	1.760	16.826	-3.518	14.918	5.426
	(6.248)	(0.086)	(0.159)	(1.984)	(4.129)	(38.857)	(8.188)	(34.384)	(12.429)
Elasticity	1.87	2.599	0.375	3.05	1.696	1.28	-1.297	1.554	6.531
Observations	99,744	99,744	99,744	99,744	99,744	99,744	99,744	99,744	99,744

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. While the sample changes in each row of models, each specification corresponds to the preferred covariate set, including demographic covariates, police department fixed effects and year by city size fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants.

Robustness Tests and Samples, Specific Crimes (Part 2)

	Violent Crime					Property Crime			
	(1) Violent Crime	(2) Murder	(3) Rape	(4) Robbery	(5) Aggravated Assault	(6) Property Crime	(7) Burglary	(8) Larceny	(9) Vehicle Theft
<i>Base Model</i>									
Police	-1.892* (0.75)	-0.007 (0.022)	-0.024 (0.086)	-0.613*** (0.186)	-1.249+ (0.651)	-9.904** (3.808)	2.194* (0.984)	-8.606** (3.107)	-3.493*** (0.674)
Elasticity	-1.28	-0.475	-0.219	-2.102	-1.176	-0.733	0.789	-0.872	-4.065
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
<i>7. Model without Apply Controls</i>									
Police	-0.650 (0.757)	-0.014 (0.024)	0.109 (0.091)	-0.505** (0.195)	-0.241 (0.66)	-8.994* (3.969)	2.953** (1.039)	-8.627** (3.275)	-3.319*** (0.691)
Elasticity	-0.44	-1.009	0.975	-1.732	-0.227	-0.665	1.061	-0.874	-3.862
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
<i>8. Illustrative Alternate IV: Award Size</i>									
Police	-0.913 (0.627)	-0.01 (0.03)	0.001 (0.08)	-0.420** (0.156)	-0.484 (0.564)	-6.194 (3.835)	1.561* (0.794)	-5.763+ (3.223)	-1.992** (0.713)
Elasticity	-0.623	-0.724	0.0113	-1.471	-0.459	-0.46	0.563	-0.586	-2.349
Observations	85,567	85,567	85,567	85,567	85,567	85,772	85,772	85,772	85,772
<i>9. Agencies that Ever Applied</i>									
Police	-1.491* (0.679)	-0.002 (0.019)	0.025 (0.076)	-0.513** (0.162)	-1.001+ (0.597)	-8.107* (3.484)	1.892* (0.886)	-6.798* (2.83)	-3.2*** (0.602)
Elasticity	-0.931	-0.102	0.211	-1.545	-0.882	-0.561	0.643	-0.645	-3.301
Observations	73,690	73,690	73,690	73,690	73,690	73,941	73,941	73,941	73,941
<i>10. Agencies that were Accepted</i>									
Police	-1.301 (0.891)	-0.008 (0.024)	0.014 (0.101)	-0.458* (0.218)	-0.849 (0.787)	-10.052* (4.590)	2.220+ (1.179)	-7.958* (3.675)	-4.314*** (0.881)
Elasticity	-0.719	-0.49	0.111	-1.135	-0.675	-0.647	0.698	-0.708	-3.876
Observations	47,396	47,396	47,396	47,396	47,396	47,548	47,548	47,548	47,548
<i>11. Agencies both Rejected and Accepted</i>									
Police	-1.158 (1.106)	-0.016 (0.03)	0.021 (0.13)	-0.358 (0.265)	-0.805 (0.993)	-9.13+ (5.547)	2.522+ (1.454)	-6.775 (4.391)	-4.877*** (1.194)
Elasticity	-0.660	-1.015	0.161	-0.932	-0.658	-0.586	0.801	-0.598	-4.432
Observations	32,091	32,091	32,091	32,091	32,091	32,179	32,179	32,179	32,179
<i>12. Excluding Withdrawn Accepted Grants</i>									
Police	-1.707* (0.74)	-0.004 (0.022)	-0.005 (0.086)	-0.551** (0.187)	-1.147+ (0.645)	-10.29** (3.832)	2.424* (0.983)	-9.143** (3.158)	-3.573*** (0.691)
Elasticity	-1.154	-0.272	-0.046	-1.891	-1.08	-0.761	0.871	-0.926	-4.158
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
<i>13. Acceptance Group by Year Fixed Effects</i>									
Police	-2.636** (0.993)	-0.006 (0.027)	-0.039 (0.113)	-0.898*** (0.259)	-1.693* (0.846)	-14.930** (5.230)	3.119* (1.302)	-12.655** (4.244)	-5.394*** (1.075)
Elasticity	-1.783	-0.417	-0.352	-3.083	-1.594	-1.105	1.121	-1.282	-6.277
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
<i>14. Application Cohort by Year Fixed Effects</i>									
Police	-2.565** (0.969)	-0.012 (0.027)	-0.066 (0.109)	-1.000*** (0.264)	-1.486+ (0.815)	-15.171** (5.196)	2.857* (1.287)	-12.250** (4.179)	-5.778*** (1.108)
Elasticity	-1.734	-0.900	-0.595	-3.433	-1.399	-1.122	1.027	-1.241	-6.723
Observations	93,081	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes and police variables are per 10,000 residents in a municipality. While the sample changes in each row of models, each specification corresponds to the preferred covariate set, including demographic covariates, police department fixed effects and year by city size fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants.

Table 2.7: Shifts in Policing Focus, Arrests for Specific Crimes

	(1) Violent Crime	(2) Murder	(3) Rape	(4) Robbery	(5) Aggravated Assault	(6) Property Crime	(7) Burglary	(8) Larceny	(9) Vehicle Theft
Arrest Rates									
Police	-0.126 (0.378)	-0.024 (0.026)	0.033 (0.044)	0.026 (0.085)	-0.161 (0.346)	6.043 (5.786)	0.197 (0.237)	3.215 (3.018)	2.631 (2.841)
Elasticity	-0.166	-1.973	0.955	0.240	-0.265	2.199	0.447	1.501	15.98
Y Mean	17.75	0.280	0.813	2.477	14.18	64.06	10.27	49.96	3.838
Observations	81,605	81,605	81,605	81,605	81,605	81,786	81,786	81,786	81,786
Number of Departments	6,411	6,411	6,411	6,411	6,411	6,410	6,410	6,410	6,410

	(10) Marijuana Sale	(11) Narcotics Sale	(12) Other Drug Sale	(13) Marijuana Possess	(14) Narcotics Possess	(15) Other Drug Possess	(16) Simple Assault	(17) DUI
Arrest Rates								
Police	0.623** (0.198)	-0.431* (0.192)	0.060 (0.192)	2.249* (0.943)	-1.302*** (0.384)	-0.437 (0.463)	-0.151 (0.864)	3.371* (1.311)
Elasticity	3.806	-3.984	0.534	1.665	-4.240	-0.915	-0.0700	1.217
Y Mean	3.815	2.526	2.631	31.50	7.163	11.15	50.37	64.55
Observations	81,423	81,423	81,423	81,423	81,423	81,423	82,438	82,438
Number of Departments	6,332	6,332	6,332	6,332	6,332	6,332	6,405	6,405

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Standard errors in each model are robust and clustered at the police department level. Crime outcomes, arrest outcomes and police variables are per 10,000 residents in a municipality. All models shown have the preferred set of covariates, including demographic covariates, application and acceptance controls for other grants, year by city size fixed effects and police department fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants. Arrest outcomes are not expressed relative to reported crime rates but as total arrests per 100,000 residents. The sample size is different for each arrest type, given the sample cleaning procedure (see Data Appendix C.2).

Chapter 3

Patrolling Public Schools:

The Impact of Funding for School Police on Student Discipline and Long-term Education Outcomes

As police officers have become increasingly common in U.S. public schools, their role in school discipline has expanded. While there is growing public debate about the consequences of police presence in schools, there is scant evidence of the impact of police on student discipline and academic outcomes. This paper provides the first causal estimate of funding for school police on student outcomes, leveraging variation in federal Community Oriented Policing Services (COPS) grants. Exploiting detailed data on over 2.5 million students in Texas, I find that funding for police in public schools results in a small but significant reduction in high school graduation and college enrollment. In contrast to higher total rates of discipline observed for low-income minority students, I find that the negative treatment effects of school police funding are more concentrated among students that are not classified as low-income and low-income White students.

3.1 Introduction

Police are an active presence in U.S. public schools. In 2014, 43% of public schools had security staff at school at least once a week, affecting over 70% of students across the country [105]. While estimates vary, government surveys suggest that there are at least 20,000 police officers working in schools [57]. These School Resource Officers (SROs) serve a number of roles, including protecting campuses from outside threats and educating students about safety and the law. However, SROs also fill another critical role: they are engaged in daily school discipline issues and administer punishments for student behavior. Despite the fact that victimizations of students have decreased in recent decades, serious school disciplinary actions have become more commonplace, increasing by nearly 200% between 2000 and 2014, and affecting Hispanic and Black students at 1.5 to 2.5 times the rate of White students [28, 105].

As police have become a fixture in public schools, policy-makers, educators, researchers, and academics are debating the merits of this approach to school discipline. Proponents of school police advocate that SROs are critical to preventing rare but catastrophic school shooting events and may also serve as positive role models to students. Critics of school police argue that SROs create a heavy-handed disciplinary culture that adversely affects learning and may further disadvantage poor and minority students in low-performing schools.

School discipline policy can have real impacts on educational at-

tainment outcomes, such as earning a high school diploma or enrolling in college. Safety is a prerequisite to learning, and policies that increase school safety and deter dangerous or disruptive behavior may have a positive effect on student academic success. Alternatively, disciplinary actions may stigmatize students and decrease their attachment to school, negatively affecting their performance. Studies in economics have found that juvenile arrests and juvenile detention decrease the probability of completing high school and increase the probability of future arrests [3, 63, 55]. The economic literature on juvenile behavioral responses to criminal sanctions has also found that juveniles may be less deterred by changes in punishment severity [68] and may be more negatively impacted by the experience of sanctions [3, 7]. By extension, citations, arrests, and referrals to juvenile detention that result from interactions with SROs may lead to future involvement with the criminal justice system, in a process often termed the "school-to-prison pipeline." Through these channels, school police can positively or negatively impact the long-term outcomes of the students that they interact with, affecting their human capital development, labor market attachment, and earnings later in life. This paper is the first study to estimate the causal impact of school police on student educational outcomes.

While there is a growing qualitative and ethnographic literature on the impact of school security on students [e.g. 86], there is little quantitative research on this topic and the causal effect of school police is unknown. Literature reviews and meta-analysis studies on school police note the lack of

rigorous empirical evidence on this topic [57, 37, 2, 11]. Studies in this space are often limited by small samples or consider simple observational pre-post or cross-sectional comparisons between schools. The best example of descriptive research in this area is [82], which uses a survey of 470 schools and a difference-in-difference design, and finds that schools that increase policing report an increase in non-serious violent crimes.

Studying the impact of school police presence on students has proved difficult for a number of reasons. First, appropriate data is hard to obtain. While schools follow a mandate to track aggregate disciplinary outcomes, detailed student-level data sets are not widely available. More importantly, information on the number of police employed in a school is not uniformly tracked because SROs are typically employed by a third party, such as a police department, rather than directly by a school district. Beyond data constraints, the assignment of police officers to particular schools is designed by school administrators, city officials, and law enforcement leaders and is non-random. Schools with higher levels of disciplinary actions, higher rates of students in poverty, higher minority populations, and lower graduation rates typically have a larger police presence. Given the selected characteristics of these school types, cross-sectional comparisons between schools that have police and those that do not will be biased. However, even when researchers examine changes in police presence in a particular school, the timing of investments in police may also be a function of changes in discipline and student behavior. If schools choose to hire police when they experience an increase in

negative student behaviors, then not only is discipline a function of policing but policing is also likely a function of discipline, and simple longitudinal or panel data analysis will be biased by this joint determination or simultaneity.

To address these measurement obstacles, I use information on federal grants received by school districts to fund police in public schools. I measure the impact of this funding on a range of student outcomes, using variation across years *within* school districts, rather than cross-sectional variation across school districts. Within a given district, I compare students enrolled in school in years when the district receives federal grant funding to students enrolled in school in years without this funding. Critically, I also account for non-random school district decisions to seek funding for police in particular years by including grant application timing as a direct control in the model.¹

This paper represents the first study to measure the causal impact of school police on student outcomes. I leverage detailed individual data on students in all public schools in Texas that enables me to track short-term changes in disciplinary actions as well as high school graduation, college attendance and graduation, and earnings and employment. The empirical design of the study exploits variation in federal Community Oriented Policing Ser-

¹The empirical approach in this study is closely related to work on the impact of COPS hiring grants for traditional police departments on municipal crime rates [103]. Likewise, this paper is also related to the larger literature on the impact of COPS grants on crime [34, 42, 106] and a growing body of economics research using quasi-experimental methods that finds that increasing police presence reduces crime rates in the general population [e.g. 25, 31, 72, 64, 29].

vices (COPS) grants that provide funding for police in public schools. Over my sample period of 1999 to 2013, over \$1 billion in federal COPS grants was distributed nationally for purposes of hiring school police officers, providing funds for school security technology, and other community partnership projects, with \$71 million awarded in Texas.

Using data on 2.5 million students in 1,167 Texas public school districts, I find that federal funding for school police modestly increases disciplinary actions for middle school students by 4% and decreases disciplinary actions for high school students by 3%. Exposure to a 3-year federal grant for school police decreases high school graduation rates by approximately 1% and decreases college enrollment rates by 3%. However, the impact of police funding differs across student race and socioeconomic status. Negative average outcomes are concentrated among students that are not classified as economically disadvantaged and economically disadvantaged White students, with long-term negative outcomes linked to differential increases in middle school discipline for these groups. Because the estimation strategy identifies the impact of marginal increases in policing, the demographic pattern of the results may be driven by the fact that poor minority students are already being disciplined at significantly higher rates than wealthier students and White students, and marginal increases in policing may allow disciplinary focus to expand to groups that do not have a high existing level of contact with school discipline.

This paper is organized as follows: Section 3.2 describes details of the federal COPS hiring grant program and student discipline in Texas, providing institutional context for the estimation strategy. Section 3.3 describes the empirical model and the data sources used in estimation. Section 3.4 discusses the results, outlining main findings and heterogeneity of the results by student demographic characteristics as well as robustness tests. Section 3.5 concludes.

3.2 Background and Institutional Context

3.2.1 Federal COPS Grants for School Police

The COPS office at the Department of Justice was originally established to fund the hiring of new police officers as part of the Violent Crime Control Act of 1994. In 1998, the COPS office extended its grant programs to include funding for police officers in schools, with the launch of pilot programs to fund police hiring in schools and partnerships between school police, schools, and other community organizations. Political interest for funding police in schools escalated after the high-profile Columbine school shooting in 1999. Policy-makers invested in school police with the intention of both preventing tragic school shootings and generally improving safety in public schools. Between 1999 and 2005, over \$750 million in funding was distributed for the purpose of hiring police officers in schools, allocating funding for over 6,500 SRO positions [22].

While Columbine inspired a surge of investment in SROs, school police and security staff did not originate in this period. The earliest school police programs began as early as the 1940s, and several large urban school districts began establishing school police programs in the 1980s and 1990s,² coinciding with a new movement of "zero-tolerance" toward student misconduct.³ Still, prior to 1999, school police were confined to large urban districts and school police programs were not widespread [2, 11]. Over the past decade and a half, new SRO programs have been founded while pre-existing SRO programs have grown; today, there are at least 20,000 SROs nationwide and over 70% of students attend schools with security staff on campus at least once a week [105, 57].

In recent years, interest in providing COPS funding for SRO programs has decreased. Federal appropriations for the school police hiring declined in the mid-2000s as part of a broader reduction in COPS funding by the Bush Administration, which had concerns about the overall effectiveness

²A sampling of the earliest recorded incidents of police operating in public schools are Los Angeles, CA (1948), Indianapolis, CA (1939), and Flint, MI (early 1950s). The National Association of School Resource Officers (NASRO) was established in 1991, also prior to the expansion of school police that was spurred by the Columbine shooting [11].

³"Zero-tolerance" policies refer to laws or school policies that require predetermined consequences for specific student offenses, without considering mitigating circumstances or context for an offense incident. These policies have expanded in scope and across geographies since the passage of the federal Gun Free Schools Act in 1994, which required automatic expulsion of any student that brought a gun on school grounds. Research on the implementation of "zero-tolerance" policies, or rules that mandate harsh discipline for a set of student infractions, is similarly sparse. An exception is a recent study by [24], which examines state-level "zero-tolerance" statutes using a difference-in-difference design and finds that these policies modestly increase overall suspension rates and have larger impacts on Black students.

in COPS grants [34]. During the Obama Administration, officials became increasingly concerned about the active role many SROs play in disciplining students, the large disparity in school discipline by student race, and the fact that student interactions with SROs may have repercussions for student involvement in the criminal justice system later in life. Given these concerns, COPS funding levels for school police have remained low since 2009. In 2014 and 2016, the Departments of Education and Justice released new guidance and resources for SRO programs, defining a narrow role for school police that excludes involvement in routine discipline and highlights the importance of disciplinary systems that do not discriminate against groups of students [33].

During the sample period of this paper, there were three broad groups of federal grants available for use in schools. The largest program, COPS in Schools (CIS), operated from 1999-2005 and provided up to \$125,000 in hiring funds per SRO over a period of 3 years. Approximately three-quarters of all COPS funding for school police has been granted through CIS. The second group of funds was administered by Secure Our Schools (SOS) grants between 2003-2012, and focused on auxiliary functions of school police. With over \$100 million distributed nationally, these flexible 2-year grants have funded security technology, security assessments, and training for school police. The final school police grants offered through the COPS office consists of school-based partnerships (SBP) with law enforcement agencies for particular projects, as well as other grant types that were granted to school district police

departments.⁴

The application process for COPS school grants is narrative-based. Each applicant is asked to describe safety problems facing their school district and their proposed approach to remedying these problems, denote any community partnerships that support the grant proposal, and state their request for assistance. Review of these applications was based on the subjective judgments of individuals at the COPS office. Given this application review process, it is likely that grant awardees were not randomly selected among school districts in each grant solicitation period. Because of this, the research design in this paper does not rely on cross-sectional variation in grant receipt between school districts, but rather focuses on within school district variation, comparing years with federal COPS funding to years that are not funded for the same school district.

3.2.2 School Resource Officers and School Discipline in Texas

The setting for this study is the state of Texas. With 5.2 million students enrolled, Texas public schools have over 10% of the U.S. student population and represent the second largest state school system after California

⁴Information on COPS grant programs for school police was obtained through the COPS website, conversations with COPS employees, and summary statistics from data on all COPS grant acceptances and rejections obtained through a Freedom of Information Act (FOIA) request for this research. An additional grant program, the Safe Schools Initiative (SSI) provided flexible funding for school and community safety and delinquency prevention, but no funds through this program were distributed in Texas during the sample period.

[84]. The student body in Texas is diverse, with minority students representing over half of the student population (see Table 3.1). Though this paper is restricted to a single state, the size and diversity of the setting make the findings informative for other contexts.

A number of advocacy and research organizations have studied student discipline in Texas, in part because of the availability of student-level discipline data in this state. Selected statistics from this research finds that nearly 6 in 10 students are suspended or expelled between the 7th and 12th grade, over 275,000 misdemeanor tickets are issued to juveniles for truancy and other student misconduct each year, and minority students and Special Education students are disproportionately disciplined relative to White students [35, 38]. In recent years, reports by the organization Texas Appleseed on student ticketing and school police in Texas prompted new legislation limiting issuance of citations and fines for misbehavior in school and mandating increased training for SROs [95, 96]. Because of a lack of comprehensive data in other states, it is difficult to know if school discipline patterns in Texas are representative of the rest of the country.

Texas has embraced the use of SROs in schools, and many communities have established "school district" police departments that are designated to serve a particular school district rather than contracting with police officers employed by local municipal police departments to serve in schools. This independent school district police model has become widespread in Texas, with

over 160 school district police departments in the state [38]. These police departments tend to be present in larger school districts and may operate in high schools, middle schools, and elementary schools. A typical police patrol ratio in a large school district is two officers per high school, one officer per middle school and rotating patrol in elementary schools. In addition to school patrol, several school district police departments in Texas have specialized units, including K-9 teams, gang suppression units, crisis response teams, traffic safety, and incident reporting hotlines.⁵ The size and budget of these police departments varies; in 2007, Houston ISD Police employed 289 staff at a cost of \$55 per student, while Edgewood ISD employed 31 staff at a cost of \$145 per student, and San Angelo ISD had a staff of 44 at a cost of \$16 per student [38].⁶

COPS grants for school police have been actively utilized in Texas, with 7% of total federal funds distributed in the state. Figure 3.1 shows that the majority of grants and funds distributed have been distributed through the CIS program, funding hiring of SROs in schools. The majority of grants were distributed between 2000 and 2004, with funding peaking at over \$16 million dollars for more than 100 new grants in 2001.⁷ The bottom panel of Figure 3.1 shows that COPS grants for school police have been consistently

⁵These characterizations of school district police departments come from web searches of police departments in Texas.

⁶Data for Houston ISD and San Angelo ISD is from the 2006-2007 school year, while data for Edgewood ISD is from the 2007-2008 school year.

⁷Throughout this paper, years refer to the academic calendar indexed by the spring semester. For example, grant statistics for 2000 cover the 1999-2000 academic year.

competitive, with grant applications outstripping grant acceptances in each year of the program.

3.3 Empirical Model and Data Sources

To assess the impact of funding for police on public school students in Texas, I exploit quasi-experimental variation in federal COPS grants for school police. My model uses panel data to measure the impact of receiving a grant within school districts over time, while also controlling for the timing of the decision to apply for COPS grants. Controlling for elective COPS grant applications addresses the non-random timing of school district decisions to expand police presence, which are likely a function of changes in school disciplinary culture and student behavior. After conditioning on the timing of a school district's decision to apply for grants, the timing of a grant acceptance is a function of the federal government's decision to fund an SRO program in a that year relative to a different year when the district applied for funding.

The primary limitation of my model is that I am unable to observe the actual employment levels of police in school districts. School districts do not directly employ police, instead they are contracted through a third party, either a municipal police department or an independent school district police department. Because I do not observe police employment, the empirical approach does not use COPS grant variation as an instrument for police presence; instead, I estimate the reduced form impact of receiving funding for

school police on student outcomes.⁸

The empirical model is as follows:

$$\begin{aligned}
ShortTermOutcome_{igdt} = & \beta_{1m}Accept_{dt} * MiddleSchool_{gt} \\
& + \beta_{2m}Apply_{dt} * MiddleSchool_{gt} \\
& + \beta_{1h}Accept_{dt} * HighSchool_{gt} \\
& + \beta_{2h}Apply_{dt} * HighSchool_{gt} \\
& + \pi X_{igdt} + \delta_t + \gamma_g + \phi_d + \varepsilon_{igdt} \\
LongTermOutcome_{idt+k} = & \alpha_1 AcceptExposure_{dt} + \alpha_2 ApplyExposure_{dt} \\
& + \tilde{\pi} X_{idt} + \tilde{\delta}_t + \tilde{\phi}_d + \nu_{idt}
\end{aligned}$$

where, i indexes students, g indexes grade, d indexes school district, and t indexes year. X_{igdt} is a vector of covariates that includes district-grade enrollment and student characteristics. At the student-level, covariates include race (Black, Hispanic, White, or other race), gender, and whether the individual is classified as a limited English proficiency (LEP) student, a Special Education student, a gifted and talented student, or an economically disadvantaged student, where economic disadvantage is an indicator for a student receiving a free or reduced lunch at school. δ_t are year fixed effects, which

⁸[103] estimates a similar model using COPS grant acceptances as an instrument for police employment in municipal districts and to determine the impact of police force expansions on local crime rates. This two-stage model is possible because municipal police departments report their employment to the Federal Bureau of Investigation each year.

capture statewide time trends in student outcomes. γ_g are grade level fixed effects, which capture differences in disciplinary actions across grades.

The student level data was obtained through the Texas Education Research Center (ERC), a research platform that allows researchers to track the state population of public school students through primary and secondary school to college and the labor market. This resource combines databases on K-12 public school students from the Texas Education Agency (TEA), post-secondary students in Texas higher education institutions from the Texas Higher Education Coordinating Board (THECB), and individuals employed in the state through the Texas Workforce Commission (TWC). I utilize information on 2.5 million students from 1,167 school districts in cohorts entering the 7th grade between 1999-2006, following students through middle school, high school and beyond graduation to college and the labor market.

The primary short-term outcomes of the analysis are whether a student received a disciplinary action and the types of disciplinary actions received, in-school suspensions, out-of-school suspensions, and expulsions. In the long-term outcome model, I focus on 7th grade student observations and measure high school graduation and college enrollment within 8 years (by age 20), college graduation, employment, and earnings within 12 years (by age 24). This approach allows me to avoid issues of attrition in the data, as students may leave school or drop out between the 7th and 12th grade. In the short-term model, I track cohorts of students beginning 7th grade in 1999-2006 through

2013, and measure disciplinary outcomes for these student groups in grades 7-12, allowing students to repeat up to two grades. Given the years spanned by the data, I am restricted to cohorts beginning the 7th grade in 1999-2006 for high school graduation and college enrollment outcomes and 1999-2001 for college graduation and labor market outcomes.

The critical variables in the short-term model are $Accept_{dt}$ and $Apply_{dt}$. These variables are constructed to match the duration of grant projects, where CIS grants last for 3 years, SOS grants last for 2 years, and other grants vary in length. In the case of a 3-year CIS grant, the variable $Apply_{dt}$ is an indicator variable for whether a grant application for a school district was submitted in year t , $t - 1$, or $t - 2$, allowing this variable to be set to 1 for the duration of the grant period in which funding would be distributed if an application was accepted. Likewise, for a 3-year CIS grant, the variable $Accept_{dt}$ is an indicator for whether a grant application covering a school district is accepted for the duration of the grant project period.⁹ For example, a school district police department that applies for and receives a CIS grant in 2000 would have both indicators set to 1 during the period 2000-2002; while if the grant application is denied, the $Apply_{dt}$ variable would be set to 1 and the $Accept_{dt}$ variable would be set to 0 for this period.¹⁰

⁹In practice, school districts do not directly apply for COPS funding. Grantees are commonly municipal police departments, independent school district police departments, or other entities. In some cases, grant applications corresponded to a geographic area that covered more than one school district. In these instances, I manually matched grants to school districts as best as possible (see Appendix C.2).

¹⁰The start time of a grant is indexed to the current academic year if a grant project (or

The grant data used in this paper was obtained through a Freedom of Information Act (FOIA) request to the COPS DOJ Office. Through this request, COPS shared information all grant acceptances and rejections since their office was founded, with information on grant program, project start date, and project end date. To identify grants for analysis in this project, I selected grants based on whether the grant program type was focused on school police or if the grant applicant had their primary jurisdiction within public schools (e.g. school district police departments).¹¹

I consider grant variables separately depending on whether the student is in middle school (7th and 8th grade) or high school (9th through 12th grade), entering these variables as interactions with school type. Given this structure, there are two acceptance variables, $Accept * MiddleSchool_{dgt}$ and $Accept * HighSchool_{dgt}$, and two comparable application variables in the model. I add this structure because in most districts students are physically separated in different school buildings across these grades and SROs likely have different capabilities and approaches to interacting with students in middle school and high school. When SRO programs are established, they typically begin operating in high schools and then expand to middle schools (and elementary schools) as they grow in size and scope, and this pattern of growth means that high

application) starts between September and March, and is indexed to the following academic year if a project starts in April through August. Throughout this project, academic years are denoted as the year of the spring semester.

¹¹The COPS office also provided information on the amount of money awarded and the number of new officers eligible for funding for accepted grants, although this information was incomplete.

schools are more likely to already have an SRO presence before they receive a grant treatment. In addition to differences in treatment across middle and high schools, students have developmental differences across these grades as well, which may impact the way that they respond to increased SRO presence.

For the long-term outcome model, the estimation approach considers future outcomes for cohorts of 7th graders and do not measure concurrent outcomes as students move through public school. In this alternate setting, the critical grant variables are defined in terms of years of exposure, as the number of years in an grant application or acceptance period divided by the student's total number of years in middle and high school. Exposure is calculated as a rate within the 6 years an "on-time" student would take to graduate high school between the 7th and 12th grades. In the 2000 CIS grant example above, a student beginning 7th grade in the year 2000 would be exposed to 3 accepted grant years and have a value of $\frac{1}{2}$ for *AcceptExposure_{dt}* and *ApplyExposure_{dt}*. These values are cumulative and account for all grant applications and acceptances that relate to a school district in the sample period, with overlapping grants only counted once. The exposure values also depend on the year that a student enters the 7th grade; in the above case, a student in the 2001 7th grade cohort would be exposed to two years of a grant and have a calculated exposure of $\frac{1}{3}$ between the 7th and 12th grades.

School districts may be covered by multiple grant applications or a grant acceptances during the sample period, and school districts can alter-

nate between states over time, switching between having an accepted grant, a rejected grant, or no application. The total reduced form impact of receiving grant funding for school police is the sum of coefficients on $Accept_{dt}$ and $Apply_{dt}$, while the impact of not receiving grant funding is the coefficient on $Apply_{dt}$. Because grant application decisions are elective choices influenced by changing the changing discipline environment and goals of school districts, $Apply_{dt}$ is a crucial control in the model, while the conditional variable $Accept_{dt}$ represents the causal reduced form impact of changes in grant funding for school police. In this framework, $Apply_{dt}$ illustrates what happens to students when a school district wants to increase their police presence but does not receive an increase in external funding to do so.¹²

The last important feature of the model are school district fixed effects, ϕ_d .¹³ These fixed effects are important because they control for unobserved differences across school districts that are constant over time. These time constant differences across districts may reflect differences in funding structures, approaches to school discipline, and school cultures, each of which are determinants of student outcomes but are unobservable in the data. Be-

¹²In their evaluation of COPS hiring grants, [34] used variation in the size of the grant award as an instrument for police force expansions rather than simply whether or not a grant was received. In this paper, it is not possible to use the size of monetary awards to measure outcomes because comparable information is not available for grants that were rejected.

¹³Students are assigned to the school district they are enrolled in during the 7th grade, rendering the output of the model "intent-to-treat" estimates. This assignment procedure assumes that students do not alter their school district in response to school police presence prior to entering the 7th grade, an assumption that is reasonable given that levels of student discipline are low in grades K-6.

cause the analysis includes school district fixed effects, the model does not make simple cross-sectional comparisons across districts that receive grants and districts that do not receive grants in a given year. Instead, the model uses variation across acceptance years, rejected years, and years with no application *within* the same school district. Appendix Figure C.1 depicts how this variation is used by comparing two hypothetical district grant histories.

The resulting identifying assumption is that conditional on the decision to apply for federal funding for school police in a given year, the timing of the acceptance of the grant proposal is not a function of changes in student outcomes within a school district. While grant variables are defined as exposure rates in the long-term model, a similar identifying assumption holds because the timing of acceptances within districts determines the total number of years of grant exposure for school district cohorts.

After accounting for a school district’s applications for grant funds, the likelihood that a district wins a grant in one year versus another application year is a function of the availability of federal funds. Given that there is a high level of variability in federal appropriations for these grants across years, the number of possible grants that can be funded in each year varies with federal interest in the grant programs, and this is a key driver of the probability that a grant application is accepted for a particular district in a specific year (Figure 3.1, bottom graph).

A common concern in studies that utilize grant variation is that

grants may be used for other purposes, or affect other aspects of a grant recipient's spending that might also affect measured outcomes. In this context, the fungibility of grant funds is not a primary concern because school districts do not directly receive grant funds. The organizations that apply for or receive COPS funding are third party police departments that have finances separate from the school districts where they work. In Texas, these third party contracting police departments are typically municipal police or designated independent school district police departments (ISD police). However, even ISD police departments that only operate within a single designated school district (discussed in Section 3.2.2) have administrations and budgets that are separate from those of school districts. This separate administration of police and school districts allays concerns that school districts themselves directly use grant funds for other school activities that could affect student educational outcomes.

3.4 Results

3.4.1 Summary Statistics

Table 3.1 provides a summary of the student data used in this project. The descriptive statistics are weighted by student-years to match the primary short-term regression analysis. The analysis covers 2.5 million students and 13.6 million student-year observations in 1,167 school districts between 1999-2013. The student sample is diverse; student observations are

42% White, 41% Hispanic, 14% Black, and 3% other race, of which 91% are Asian. 49% of the sample is categorized as economically disadvantaged, or low-income, a designation that is derived from whether a student receives a free or reduced price lunch at school.

Disciplinary actions are common with 27% of students receiving disciplinary actions each year. On average, 22% of students receive an in-school suspension each year and 10% receive out-of-school suspensions, while less than 1% are expelled.¹⁴ Over the long-term, 70% of 7th grade cohorts graduate from a public high school in Texas, and 47% enroll in college within 7 years following the 7th grade.¹⁵ 11 years after the 7th grade (or 6 years after 12th grade if a student graduates high school "on-time"), 19% of 7th grade cohorts in the sample graduate from college and 61% are employed in Texas.

¹⁴In the ERC data, not all disciplinary actions are coded as suspensions or expulsions, and may include removals from class or another sanction that is not specified. Additionally, students may receive more than one discipline type in a given year (e.g. an in-school suspension and an out-of-school suspension), allowing the sum of the components not to equal the total rate of disciplinary actions.

¹⁵The high school graduation rates shown in Table 3.1 are lower than the longitudinal graduation rates reported by the state of Texas, which were 89% in 2014-2015. These rates are calculated for different populations; official statistics consider 4-year graduation rates for cohorts of 9th graders, while I consider graduation rates within 8 years for 7th graders. While the rates I calculate are lower than the state's, they are actually more liberally defined; I allow students two additional years to complete high school and permit them to transfer to and graduate from any public school district in the state. However, the key difference is that official statistics remove students from their calculation if they leave school but do not dropout, removing these students from the denominator of the rate. Students may be classified as "leavers" if they leave to attend a private school or a school in another state, begin homeschooling, leave for health reasons, or leave for other unverified reasons. Recent research shows that school districts have significant latitude when defining graduation rates and frequently fail to appropriately account for student attrition [e.g. 81].

COPS grants for police in schools affect a large portion of student-years in the data, with 41% of observations corresponding to a grant application year and 22% of observations corresponding to a grant acceptance year (Table 3.1). These statistics imply a student-weighted grant acceptance rate of 54%. Table 3.2 shows statistics on the COPS grants used in analysis, also weighted by student-years to match the sample used in the analysis. Over the time frame of the study, there were 1,066 applications for COPS grants that fund school police and 486 grant acceptances.¹⁶ Awarded grants designate funding for 2.25 eligible SROs per school district on average, with total funds of \$287,102 per school district (weighted by student-years). On a student-weighted basis, 80% of students attended school districts that applied for an average of 3 grants, while 68% of students attended school districts that received a grant in the sample period.¹⁷

¹⁶These counts are calculated to reflect counts of new grant-years by school districts. They do not match the underlying grant data given that some grants cover multiple school districts and school districts with 2 new grants in a single year are only counted once. The percentage distribution by grant type does allow for double counting however, reflecting the fact that school districts may be covered by more than one application or acceptance in a given year.

¹⁷When grant characteristics are not weighted by student observations, 41% of districts applied for an average of 1 grant during the sample, 24% of all districts were ever accepted for a grant, and 34% of districts were ever rejected for a grant. These numbers are lower than the weighted characteristics because larger school districts were more likely to apply for and receive COPS funding.

3.4.2 Baseline Results

In the baseline analysis, I estimate the models described in Section 3.3. These models consider the impact of the treatment to be uniform across all student groups, though effects are allowed to vary across middle schools and high schools.

Table 3.3 displays the baseline results for the short-term discipline outcomes. The grant acceptance coefficients show that, conditional on applying for a grant, receiving a grant for police decreases disciplinary actions for high school students and increases disciplinary actions for middle school students. For high school students, COPS grants decrease in-school suspensions by 5%, and expulsions by 22% relative to the sample mean.¹⁸ For middle school students, the effect of a grant acceptance conditional on application is a 4% increase in total disciplinary actions and a 8% increase in out-of-school suspensions. A depiction of the impact of a grant application and acceptance on disciplinary actions by grade is shown in Appendix Figure C.2.

Table 3.4 measures the impact of grant funding on a more detailed set of short-term student outcomes. This table shows that the increase in disciplinary actions for middle school students is primarily driven by a 6% increase in school conduct code violations. Middle school students also expe-

¹⁸Throughout this paper, percentage effects are calculated relative to the outcome mean for the entire sample period, rather than a "pre-treatment" period because districts may be treated multiple times within the sample period.

rience a 10% increase in referrals to Disciplinary Alternative Education Programs (DAEP), or designated learning centers for students with more than 3 days of an out-of-school suspension. In contrast, disciplinary reductions may be related to more serious offenses, including declines in substance abuse and sexual conduct, as well as a 14% decline in referrals to juvenile detention or Juvenile Justice Alternative Education Programs (JJAEP).

The coefficients on $Apply_{dt}$ in Tables 3.3 and 3.4 often oppose the sign of the coefficients on $Accept_{dt}$. An example is the impact on high school expulsions, which shows an increase in expulsions for students in districts that applied for grants and an equivalent decrease in districts that received an acceptance conditional on application. In effect, these coefficients imply that a district with an accepted grant has a total impact of zero on expulsions, but that expulsions would have increased if funding were not received. The application coefficients imply that the timing of applications may be related to anticipated increases in punishable student behaviors or a change in a district's approach to school discipline. This control is critical to estimating the unbiased effect of an increase in school police funding, as measured by the $Accept_{dt}$ variables.

Federal police grants also have a material impact on student attendance and enrollment decisions. Table 3.4 shows that grant receipt increases the likelihood that middle school students exit school and fail to enroll in a Texas public school (in any district in the state) in the following year by 8%

relative to the sample mean. This exit effect has the same magnitude as a decline in students repeating middle school grades.¹⁹ For high school students, the expansion of funding for SROs results in a 10% increase relative to the sample mean.

Table 3.5 provides the results of exposure to COPS grants on the long-term outcomes of cohorts of 7th graders. As discussed above, these variables define the proportion of years that a 7th grader is in a district that applied for grant funding and the proportion of years that he/she is in a district that received a grant.²⁰ The long-term analysis shows that COPS grant funding for school police significantly reduces high school graduation and enrollment in 2-year colleges. If students are exposed to one 3-year CIS grant, the coefficient magnitudes imply that high school graduation rates will decline by 1% and enrollment in 2-year colleges will decline by 5%. These largely negative long-term outcomes counter the short-term discipline gains for high school students displayed in Table 3.3.

¹⁹Given that SRO funding appears to impact student attrition, it is important to understand whether selection in student exits is driving the declines in school discipline in high school. As part of the analysis for this project, I conducted a number of tests to understand this relationship, including investigating whether exits are more likely for students with a history of disciplinary actions as well as restricting the sample to students that remain in school until the 11th grade. I find limited evidence that selection is driving the decline in discipline that is observed for high school students. These results are available upon request.

²⁰As noted above, school district assignments throughout the analysis are linked to school district enrollment in the 7th grade, even if a student transfers to another district. High school graduation is defined to allow a student to graduate from any public high school in the state. This school district assignment procedure renders the grant impact as "intent-to-treat" estimates.

As discussed above, the primary limitation to this study is that I am unable to observe the number of police officers working in schools, and therefore, I cannot estimate a first stage impact of federal grant receipt on school police presence. While SROs are contracted to work in schools through outside police departments, school district budget information can provide suggestive evidence of changes that result from COPS grants. In Appendix Table C.1, I estimate the impact of COPS grant variables on various school district security budget measures, in the same student-weighted regression framework as the general results. This exercise shows that school districts that are affected by COPS grants for school police are significantly more likely to list a security fund line item in their budget records (column 4), while there are no statistically meaningful impacts on security expense levels or the ratio of security spending to the total operating budget. This pattern is not surprising, given that school districts do not directly receive grants for school police, but grants are instead administered to law enforcement agencies that contract with schools. Nevertheless, the increase in school districts recording security expenses when a grant is received suggests that the grants have a meaningful impact on police and security presence in schools.

3.4.3 Treatment Heterogeneity

3.4.3.1 Heterogeneous Effects across Race and Income

The baseline results present a puzzle: how is it that increases in funding for police in public schools reduce high school disciplinary actions and also reduce ultimate high school graduation and college enrollment? A closer examination of differential impacts by student demographic characteristics helps explain this seeming contradiction in the baseline models.

In Tables 3.6, 3.7, and 3.8, I consider treatment effects for different student demographic groups, split by race and socioeconomic status, using the economic disadvantage indicator for students that receive a free or reduced lunch. I consider differences by race because of the large disciplinary gaps that have been documented by researchers, policy-makers, and advocates. Similarly, I consider student poverty because poorer students also experience higher rates of discipline and are less likely to graduate from high school and attend college.²¹

The demographic treatment effects in these tables are striking. Students *not classified as economically disadvantaged* experience substantial

²¹I have also examined how the results vary by characteristics of school districts. This exercise shows that treatment effects are clustered in school districts that have high levels of disciplinary actions, low graduation rates, and a high proportion of non-white students, but are also present in school districts that have high and low levels of economically disadvantaged students. Within the school district categories that show effects of grant treatment, the demographic patterns are generally comparable to those in the broader sample. The results are omitted due to space constraints but are available on request.

increases in disciplinary actions across race groups, while disciplinary actions also increase for White students and other race students that are economically disadvantaged. For these groups, the treatment coefficients on disciplinary actions are 1.3 to 3.1 times the size of the average coefficient for the total middle school population of 0.01 shown in Table 3.3. In high school, students that *are economically disadvantaged* are those that experience declines in disciplinary actions due to increases in COPS funding, with the largest effects for Black and Hispanic students. With the exception of expulsions, students that are not economically disadvantaged do not experience the same reductions in disciplinary actions in high school grades from a grant treatment.

The demographic treatment effects starkly differ from the large total gaps in discipline by race and socioeconomic class. In Appendix Tables C.4, C.5, and C.6, I display expanded output of the model that includes the estimated baseline "main" effects for demographic groups. Separate from the treatment impacts related to funding for police, the baseline probability that students categorized as economically disadvantaged receive a disciplinary action is 4 to 12 percentage points higher than students that are not economically disadvantaged within each race. Across races, the main effects paint a stark picture; relative to White students, the probability that Black and Hispanic students receive a disciplinary action is 10-11 and 2-7 percentage points higher, respectively. Disciplinary differences by race are more pronounced for

students that are not classified as economically disadvantaged.²²

The demographic treatment patterns in Table 3.6 are echoed in the more detailed set of short-term student outcomes. Table 3.7 shows that when school districts receive funding for police, conduct code violations increase for middle school students that are not economically disadvantaged as well as White students that are economically disadvantaged. Likewise, these student groups exhibit an increase in days absent from school and referrals to Disciplinary Alternative Education Program (DAEP) due to the treatment. In contrast, high school treatment effects for students that are economically disadvantaged show larger declines for conduct code violations, as well as more serious violent, and substance abuse offenses.

Turning to the long-term student outcomes, treatment effects for different demographic groups largely align with the short-term discipline results. The relatively wealthier students that had increases in middle school disciplinary actions when school districts received police funding likewise have negative long-term treatment outcomes. These negative treatment impacts span high school graduation, college enrollment and college graduation. These negative effects are strongest for Black students that are not economically disadvantaged. For these students, exposure to a 3 year grant for police funding

²²The omitted reference category in these models are other race students that are not economically disadvantaged. Each model also includes controls for special education, LEP, and gifted and talented status (output suppressed). Special education and LEP students are significantly more likely to be disciplined, while gifted and talented students are less likely to be disciplined.

reduces the likelihood of college enrollment by 3%, college graduation by 4%, employment by 2%, and income by 3%. Higher income white students experience a reduction in enrollment in 4 year colleges, while higher income Hispanic students show a reduction in enrollment in 2 year colleges.

Likewise, low-income White students, who also experienced increases in disciplinary actions in middle school, show significant reductions in long-term outcomes. This group drives the reduction in high school graduation rates observed in the total sample in Table 3.5, with the impact of exposure to a 3 year grant reducing high school graduation rates by 4% for this group. Low-income whites also experience significant reductions in 2-year college enrollment as a result of increased funding for school police.

The long-term results for economically disadvantaged Black and Hispanic students, who experienced general reductions in high school discipline, are mixed. Low-income Black students have no treatment effect on high school graduation rates, and a reduction in 2-year college enrollment rates, but also experience a slight increase in 2-year college graduation rates. Economically disadvantaged Hispanic students show the most positive treatment effects, with exposure to a 3 year grant increasing in college graduation by 7%, employment by 1.5%, and income by 3% for this group.

It is not surprising that funding for police in public schools differentially impacts students with different demographic characteristics, but the observed patterns contrast with *a priori* concerns that SROs disproportionately

disadvantage poor minority students. Given that total school discipline levels are markedly higher for poor and minority students, it is possible that when schools receive a spike in funds for school police programs, SROs expand their focus to disciplining students that have relatively lower levels of infractions but may not show better school behavior. At the same time, the suggestive positive treatment effects observed for students that are already disciplined the most may be a result of constructive deterrence. Overall, the demographic analysis implies that a student's experience with school discipline at an early age has ramifications for their academic success in high school, post-secondary educational attainment, and earnings. Negative school discipline experiences may shape the way that students are perceived by teachers, school administrators, and peers, and may also shape a student's confidence and attachment to school. Though this study provides insight into the differential treatment effects of funding for police in public schools, it is difficult to determine the precise mechanisms driving these patterns given limitations in the data.²³

²³I have also split these groups by gender, LEP status, and special education status. Consistent with the other results in this paper, the treatment effects for these groups do not always align with the baseline differences in main effects. In the gender analysis, the probability that male students are disciplined is 6-13 percentage points higher than the probability that female students are disciplined. I find that within race and income groups, male and female students have roughly similar total treatment effects on disciplinary actions, with the exception of larger increases in disciplinary actions for economically disadvantaged White females. Economically disadvantaged White female students have more negative outcomes than their male counterparts, and appear to drive the reductions in high school graduation rates observed in the average treatment coefficient. In the LEP and special education analysis, LEP and special education students have worse baseline outcomes and are more likely to be disciplined than other students within race and income groups. Non-LEP students that are not economically disadvantaged have higher disciplinary increases in middle school than LEP students and these treatment effects translate to worse long-

3.4.4 Robustness Tests of the Baseline Model

In this section, I conduct several robustness tests of the baseline model. First, I test the validity of the identification assumption, that conditional on grant application decisions, the timing of COPS grant acceptance is not a function of changes in student outcomes within a school district. Figure 3.2 interacts the accept and application variables with year indicators before and after treatment, to examine how treatments are related to changes in disciplinary outcomes over time. Each of the graphs shows different coefficients from the same regression, plotting differential effects for middle school and high school students.²⁴

These graphs tell an interesting story. On the left side of Figure 3.2, it is clear that the timing of COPS grant acceptances are unrelated to pre-treatment changes in student disciplinary actions for middle school or high school students. The top left graph shows an increase in aggregate disciplinary actions for middle school students in the post treatment period, while

term outcomes. Middle school discipline treatment effects are larger for low-income White students with special education status, and this treatment gap also results in worse long-term outcomes for this group. These results are omitted but are available on request.

²⁴Due to computational constraints given the size of my data set, these graphs were produced using data collapsed to the school district-grade-year level and weighted by the number of students within these cells. Because a school district may receive multiple grants within a sample period, these graphs are created by duplicating the data for each possible treatment year and stacking these data sets to form a "pseudo panel." In each sub-panel of the analysis, only the designated treatment year is considered over time, and other concurrent treatments in adjacent years are included as model controls. In this analysis I include treatments in 2000-2013, so that each centered treatment year has at least one year of observed pre-treatment data.

the bottom left graph shows a decline in disciplinary actions for high school students following the treatment.

When grant application variables are displayed over time, interesting relationships emerge before and after treatment (right side of of Figure 3.2). Decisions to apply for school police funding are made following increases in disciplinary actions for middle school students. School districts appear to be interested in receiving funding for police when they experience higher levels of negative student behaviors, or have increased the punitiveness of their school discipline enforcement. This pre-application increase in disciplinary actions is not observed for high school students. Consistent with the baseline model results, school districts also have modest increases in disciplinary actions when they apply for but do not receive federal funding for school police. These patterns underscore the importance of conditioning the results on the timing of application decisions, as these decisions are non-random and are related to the outcomes of interest. By directly controlling for grant applications, I am able to consider the causal impact of funding for school police, conditional on a district's choice to seek this funding. Appendix Figures C.3 and C.4 displays comparable graphs with similar treatment patterns for in-school suspensions, out-of-school suspensions, and expulsions.

To examine potential mechanisms, I explore differential treatment effects by grant type and grantee type. Appendix Table C.2 shows that COPS in Schools (CIS) grants drive both the short-term and long-term results on

student outcomes. Because CIS grants fund police hiring rather than school police technology, equipment or other initiatives, this analysis suggests that changes in the level of public school police presence affect students more than funding for other school police functions.

Next, I split school districts by police department type: police departments designated to operate in independent school districts (ISD Police) or other police department types (Not ISD Police) (Appendix Table C.3). The increases in discipline for middle school students arise from districts that do not have ISD police departments, and consequently, negative long-term student outcomes are concentrated in these districts as well. ISD police are likely to already have a substantial presence in the school districts that they serve, so grants may cause a larger proportional change in school police presence from other police department types (non-ISD police), resulting in larger negative student impacts for these police department types. Non-ISD police departments may also be more dependent on federal funding to supply officers to school districts, meaning that changes in grant funding have more of a direct impact on the school districts they serve. Alternatively, designated ISD police may be better equipped to operate in schools relative to non-ISD police because they may have a greater level of experience working in schools, they may differ in management structure and focus, or they may have better training opportunities that prepare them to work in school environments.

In additional tests, I consider sample restrictions based on the grant

histories of school districts. Tables 3.9 and 3.10 show that the results are similar when the sample is restricted to districts that ever applied for grants, ever received a grant, and were both accepted and rejected for grants (Specifications 2-4). In these specifications, the negative impact of grant exposure on high school graduation and college enrollment is slightly smaller than in the baseline model and not always significant, possibly because of the smaller sample size in these models. In Specification 5, I group districts into four categories of districts that have only been accepted for grants, only been rejected for grants, both been accepted and rejected, and never applied, and then include separate year time effects for each of these groups. While this specification allows districts with different grant history types to have differential time trends, the estimates are similar to the baseline specification.

Finally, I include regressions that collapse the data to the district-year level and are not weighted by the number of students in a district in Specification 6. The short-term regressions do not show similar treatment patterns to the baseline model, with significant reductions in out-of-school suspensions in middle school. Likewise, the long-term model shows weaker effects than the baseline. These estimates are consistent with the finding that the treatment effects are driven by larger school districts.

3.5 Conclusion

The widespread use of police officers in public schools is relatively recent development. While school police programs have gained popularity as a policy to protect students against rare but tragic school shooting events, in practice, these officers are actively involved in the enforcement of school discipline. When school police officers are involved in the daily lives of students, they have the potential to alter student behavior, disciplinary consequences, attachment to school, and long-term educational attainment. Though the potential consequences of school police interventions are large, there is little evidence of their effectiveness.

This study provides the first causal estimate of funding for police in public schools on students. Using variation in federal COPS grants, I measure the impact of receiving an increase in grant funding for school police, conditional on decisions to apply for this federal funding. By comparing student outcomes within school districts, I estimate the impact of grant receipt on students exposed to increases in school police funding relative to students that were not exposed to this treatment within the same district. This strategy addresses biases related to both the non-random assignment of police to particular school districts and the non-random timing of investments in police within school districts.

Using detailed data on 2.5 million public school students in the state of Texas, I find that increasing funding for school police increases disci-

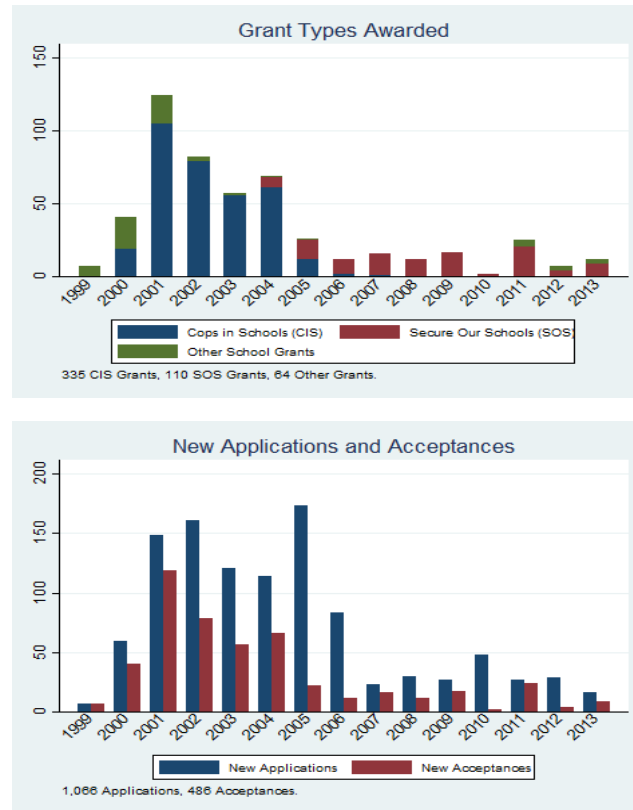
plinary actions for middle school students and decreases disciplinary actions for high school students. Over the long-term, exposure to federal funding for school police results in small but significant declines in high school graduation rates and college enrollment.

However, the results of increasing funding for school police vary substantially across student demographic groups. I find that expansions in funding for school police increase disciplinary actions for student groups that have lower baseline levels of involvement with school discipline. Namely, grants for school police increase middle school disciplinary actions for students that are relatively higher income, as well as low-income White students. At the same time, low-income Black and Hispanic students experience a reduction in high school disciplinary involvement as a result of grant funding. Over the long-term, outcomes across students reflect this pattern. While long-term educational attainment treatment effects are mixed for low-income Black and Hispanic students, these treatment effects are largely negative for relatively higher income students and low-income White students.

More research is needed to understand the ways that police affect students, as well as how the utilization of public school police compares to other approaches to school discipline. Future work should evaluate best practices in school discipline as well as the cost-effectiveness of different disciplinary approaches.

3.6 Tables and Figures

Figure 3.1: COPS Grant Funding for School Police in Texas



All grant tabulations above are calculated at the district level and are not weighted by the number of students in a district, as is done in the remainder of the analysis. Grants are attributed to the year in which they were awarded and different grant types last for different project lengths. Hiring grants are attributed to the years in which they were applied for or awarded. Grant tabulations are conducted at the school district level, rather than the grant level, to match the variation used in analysis. This means that grants awarded to multiple districts are counted more than once, and multiple grants awarded to a particular district in a given year are collapsed. Coded years in these graphs correspond to the spring of an academic year; for example, the 2000 grant tally covers the 1999-2000 academic year.

Table 3.1: Summary Statistics

	Mean	S.D.
Number of Districts	1,167	
Number of Students	2,515,026	
Number of Student-Years	13,596,579	
<u>Demographic Controls</u>		
Number of Students in Grade	2,386	(3,085)
% Male	0.51	(0.50)
% White	0.42	(0.49)
% Black	0.14	(0.35)
% Hispanic	0.41	(0.49)
% Limited English Proficiency	0.06	(0.24)
% Special Education	0.13	(0.34)
% Gifted	0.11	(0.32)
% Economically Disadvantaged	0.49	(0.50)
<u>Grant Variables</u>		
Acceptance-Year	0.29	(0.45)
Application-Year	0.48	(0.50)
<u>High School Outcomes</u>		
Any Disciplinary Action	0.27	(0.44)
Suspension (In-School)	0.22	(0.42)
Suspension (Out-of-School)	0.10	(0.30)
Expulsion	0.004	(0.06)
Low Level Offense	0.19	(0.40)
High Level Offense	0.06	(0.24)
<u>Long-term Outcomes (9th Grade Cohorts Only)</u>		
Graduate	0.70	(0.46)
Enroll in College within 2 years after HS	0.47	(0.50)
Enroll in 2-year College within 2 years after HS	0.38	(0.48)
Enroll in 4-year College within 2 years after HS	0.21	(0.41)
Graduated College within 6 years after HS	0.19	(0.39)
Graduated 2-year College within 6 years after HS	0.06	(0.24)
Graduated 4-year College within 6 years after HS	0.14	(0.35)
Employed 6 years after HS	0.61	(0.49)
Income 6 years after HS (\$)	12,993	(17,954)
Income Conditional on Employment 6 years after HS (\$)	21,374	(18,739)

This table displays summary statistics of covariates and outcomes weighted by student-years. Long-term outcomes are calculated only for 7th grade cohorts and are weighted by the number of students in these cohorts.

Table 3.2: Summary Statistics, Grant Data Weighted by Student Observations

	Acceptance		Application	
<i>(Variables weighted by number of students in a district)</i>	Mean	S.D.	Mean	S.D.
Number of New Application/Acceptance-Years	486		1,066	
Cops in Schools Grant (CIS)	0.72	(0.45)	0.77	(0.42)
Secure Our Schools Grant (SOS)	0.22	(0.41)	0.15	(0.36)
Other School Grant	0.14	(0.35)	0.14	(0.35)
Grant Applying Group is a School District Police Department	0.40	(0.49)	0.41	(0.49)
Number of School Districts Per Grant	5.02	(5.63)	5.19	(6.24)
Eligible Officers per School District	2.25	(3.65)		
Total Award per School District (\$)	287,102	(411,752)		
Total Award per Student (\$)	12.15	(30.12)		
<i>(Variables weighted by number of students in a district)</i>	Mean	S.D.		
Number of Districts	1,167			
Ever Accepted	0.68	(0.47)		
Ever Rejected	0.67	(0.47)		
Ever Applied	0.80	(0.40)		
Both Accepted and Rejected	0.56	(0.50)		
Number of Acceptances	1.71	(2.15)		
Number of Rejections	1.39	(1.30)		
Number of Applications	3.10	(2.77)		

This table shows summary statistics for the COPS grants measured in this project. Counts of new grant-years and districts are not weighted, while the values of grant characteristics are weighted. Grant counts are calculated at the school district-grant-year level, collapsing multiple acceptances and applications for a single district in a given year, and duplicating counts for grants that cover multiple school districts. The percentages of grant type are calculated as portions of the acceptance count and the application count and are weighted by student observations. These grant type rates account for multiple grants for a single district in a given year, allowing their sum to exceed 100%. Because values other than grant and district counts are weighted by student observations, they represent the proportion of students attending districts with certain grant characteristics, or the weighted average of a value.

Table 3.3: Short-term Student Discipline Outcomes

	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
Acceptance Effects				
Middle School: Accept	0.010* (0.004)	0.005 (0.004)	0.009* (0.004)	-0.0001 (0.0003)
High School: Accept	-0.007+ (0.004)	-0.010* (0.005)	-0.002 (0.003)	-0.0009*** (0.0003)
Application Effects				
Middle School: Apply	0.002 (0.004)	-0.001 (0.004)	0.004 (0.004)	-0.0002 (0.0003)
High School: Apply	0.002 (0.004)	0.003 (0.003)	-0.003 (0.002)	0.0009*** (0.0002)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577
Accept Mean	0.286	0.286	0.286	0.286
Y Mean: Middle School	0.281	0.239	0.115	0.004
Y Mean: High School	0.260	0.212	0.093	0.004
% Effect of Conditional Grant Receipt: Middle School	3.7%	1.9%	7.5%	-3.2%
% Effect of Conditional Grant Receipt: High School	-2.6%	-4.8%	-2.6%	-22.4%

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Robust standard errors clustered at the school district level are shown in parentheses. All models include a full set of covariates for student population in a school district-grade, student gender, race (White, Black, Hispanic, or Other Race), economic disadvantage status, limited english proficiency (LEP) status, special education status, gifted and talented status. Models also include grade fixed effects, year fixed effects, and school district fixed effects. This table shows concurrent impacts of a COPS grant application and grant acceptance on student discipline. Apply variables are coded as indicators of the period when a grant project would be active if a grant is approved. Likewise, Accept variables are indicators for receipt of for the duration of a grant project period. Middle school effects are for 7th and 8th grade, while high school effects are for grades 9-12. Means shown are for the entire sample period.

Table 3.4: Expanded Short-term Student Outcomes

	(1) Conduct Code Violation	(2) Violent Offense	(3) Substance Abuse Offense	(4) Sexual Conduct Offense	(5) Weapon Offense	(6) DAEP	(7) JJAEP	(8) Absent Days	(9) Repeat Grade, Next Year	(10) School Transfer, Next Year	(11) District Transfer, Next Year	(12) Don't Enroll, Next Year
Acceptance Effects												
<i>Middle School: Accept</i>	0.014** (0.004)	-0.0004 (0.003)	0.0006 (0.0004)	0.0001+ (0.00006)	-0.0002* (0.0001)	0.005+ (0.003)	0.0002 (0.0002)	0.142 (0.094)	-0.005* (0.002)	0.004 (0.007)	-0.0001 (0.0012)	0.004* (0.002)
<i>High School: Accept</i>	-0.003 (0.004)	0.0005 (0.0008)	-0.0006+ (0.0004)	-0.00008* (0.00004)	-0.0003* (0.0001)	0.002 (0.002)	-0.0004* (0.0002)	-0.089 (0.108)	0.007* (0.003)	-0.008 (0.006)	0.0008 (0.0008)	-0.002 (0.002)
Application Effects												
<i>Middle School: Apply</i>	0.003 (0.005)	0.002 (0.002)	-0.0006 (0.0004)	-0.00005 (0.00008)	0.0002* (0.0001)	-0.002 (0.002)	-0.0005* (0.0002)	-0.373*** (0.097)	-0.006* (0.003)	0.002 (0.012)	-0.003* (0.0015)	-0.004* (0.002)
<i>High School: Apply</i>	-0.002 (0.004)	-0.0008 (0.0006)	0.0005 (0.0003)	0.00007 (0.00004)	0.0001 (0.0001)	-0.003+ (0.002)	0.0003+ (0.0001)	0.185 (0.130)	0.002 (0.003)	-0.004 (0.006)	0.0017* (0.0008)	0.003+ (0.002)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,292,626	11,684,325	11,684,325	11,684,325	11,684,325
Accept Mean	0.286	0.286	0.286	0.286	0.286	0.286	0.286	0.291	0.306	0.306	0.306	0.306
Y Mean: Middle School	0.248	0.022	0.009	0.0005	0.0009	0.044	0.0025	7.62	0.018	0.504	0.078	0.051
Y Mean: High School	0.236	0.018	0.016	0.0004	0.0008	0.043	0.0026	10.21	0.077	0.127	0.061	0.083
% Effect of Conditional Grant	5.5%	-2.0%	6.6%	21.1%	-25.9%	10.2%	6.0%	1.9%	-30.5%	0.7%	-0.2%	7.9%
Receipt: Middle School	-1.1%	2.9%	-4.0%	-21.2%	-37.0%	4.3%	-14.3%	-0.9%	9.5%	-6.4%	1.2%	-2.5%
Receipt: High School												

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Robust standard errors clustered at the school district level are shown in parentheses. All models include a full set of covariates for student population in a school district-grade, student gender, race (White, Black, Hispanic, or Other Race), economic disadvantage status, limited english proficiency status, special education status, gifted and talented status. Models also include grade fixed effects, year fixed effects, and school district fixed effects. This table shows concurrent impacts of a COPS grant application and grant acceptance on student discipline. Apply variables are coded as indicators of the period when a grant project would be active if a grant is approved. Likewise, Accept variables are indicators for receipt of for the duration of a grant project period. Middle school effects are for 7th and 8th grade, while high school effects are for grades 9-12. Means shown are for the entire sample period. The DAEP outcome is referral to a Discipline Alternative Education Program, or a learning center for students with more than 3 days of an out-of-school suspension. The JJAEP outcome is referral to a Juvenile Justice Alternative Education Program, or a juvenile detention center school.

Table 3.5: Long-term Student Outcomes

	(1) Graduate High School	(2) Enroll College	(3) Enroll: 2-year College	(4) Enroll: 4-year College	(5) Graduate College	(6) Graduate: 2-year College	(7) Graduate: 4-year College	(8) Employed	(9) Income	(10) Income Given Employment
Acceptance Effect										
Accept Exposure	-0.017* (0.007)	-0.029** (0.011)	-0.037** (0.011)	-0.001 (0.005)	-0.0001 (0.004)	-0.0007 (0.002)	-0.002 (0.005)	0.001 (0.006)	355.1 (295.0)	558.4 (406.1)
Application Effect										
Apply Exposure	0.004 (0.006)	0.004 (0.007)	0.008 (0.006)	-0.002 (0.004)	-0.002 (0.004)	-0.002 (0.001)	0.0005 (0.003)	0.002 (0.004)	68.0 (115.4)	70.4 (148.7)
Student-Year Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455
Accept Mean	0.698	0.471	0.376	0.211	0.187	0.0619	0.139	0.608	12,993	21,374
Y Mean	0.255	0.255	0.255	0.255	0.401	0.401	0.401	0.401	0.401	0.388
% Effect of Conditional Grant Exposure	-2.5%	-6.2%	-9.7%	-0.6%	-0.1%	-1.1%	-1.2%	0.2%	2.7%	2.6%

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Robust standard errors clustered at the school district level are shown in parentheses. Long-term models include 7th grade cohorts and track their outcomes in future years. High school graduation is defined as graduating from any Texas public high school within 8 years after the 7th grade, college enrollment is measured as enrolling in any Texas college (private and public) within 8 years after the 7th grade. College graduation is similarly defined and is measured within 11 years of the 7th grade. Employment and earnings within Texas are measured as outcomes in the 11th year after the 7th grade. Students without employment records or with zero earnings are included in these outcomes with zero values. All models include a full set of covariates for student population in a school district-grade, student gender, race (White, Black, Hispanic, or Other Race), economic disadvantage status, LEP status, special education status, gifted and talented status. Models also include grade fixed effects, year fixed effects, and school district fixed effects. This table shows the impact of exposure to a COPS grant acceptance or application, defined as the number of years in either of these states divided by 6 (the number of years an "on-time" student is enrolled between 7th and 12th grade).

Table 3.6: Short-term Student Discipline Outcomes, by Demographic Group

	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
Middle School - Accept Effects				
Economic Disadvantage				
Black	-0.001 (0.009)	-0.006 (0.014)	-0.002 (0.009)	-0.0013 (0.0009)
Hispanic	-0.006 (0.008)	-0.009 (0.008)	-0.004 (0.005)	-0.0005 (0.0006)
White	0.013* (0.006)	0.005 (0.006)	0.016*** (0.005)	-0.0003 (0.0005)
Other Race	0.031** (0.011)	0.026* (0.011)	0.018** (0.006)	0.0003 (0.0006)
No Economic Disadvantage				
Black	0.031*** (0.008)	0.024** (0.009)	0.018** (0.007)	-0.0001 (0.0006)
Hispanic	0.022*** (0.006)	0.014* (0.006)	0.016** (0.005)	0.0006+ (0.0003)
White	0.017*** (0.004)	0.011* (0.005)	0.011*** (0.003)	0.0002 (0.0002)
Other Race	0.024*** (0.007)	0.017* (0.007)	0.014*** (0.004)	0.0003 (0.0004)
High School - Accept Effects				
Economic Disadvantage				
Black	-0.017** (0.005)	-0.021** (0.007)	-0.016** (0.005)	-0.0017** (0.0006)
Hispanic	-0.017** (0.006)	-0.018** (0.007)	-0.003 (0.004)	-0.0008 (0.0005)
White	0.0002 (0.005)	-0.002 (0.006)	0.004 (0.003)	-0.0008+ (0.0005)
Other Race	0.011 (0.010)	0.005 (0.010)	0.012+ (0.006)	-0.0009 (0.0006)
No Economic Disadvantage				
Black	-0.002 (0.009)	-0.007 (0.008)	-0.0001 (0.006)	-0.0008* (0.0004)
Hispanic	0.004 (0.008)	-0.0001 (0.008)	0.003 (0.004)	-0.0012*** (0.0004)
White	0.002 (0.004)	-0.003 (0.005)	0.001 (0.002)	-0.0008*** (0.0002)
Other Race	-0.001 (0.007)	-0.005 (0.006)	0.003 (0.004)	-0.0004 (0.0004)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Table 3.3. While not shown due to space constraints, all models include application variables interacted with the same set of demographic categories as shown above. Demographic interactions in this table are non-additive and represent mutually exclusive student groups.

Table 3.7: Expanded Short-term Student Outcomes, by Demographic Group

	(1) Conduct Code Violation	(2) Violent Offense	(3) Substance Abuse Offense	(4) Sexual Conduct Offense	(5) Weapon Offense	(6) DAEP	(7) JJAEP	(8) Absent Days	(9) Repeat Grade, Next Year	(10) School Transfer, Next Year	(11) District Transfer, Next Year	(12) Don't Enroll, Next Year
Middle School - Accept Effects												
<i>Economic Disadvantage</i>												
Black	0.006 (0.008)	-0.018* (0.008)	-0.001+ (0.001)	0.0004+ (0.0002)	-0.0008** (0.0003)	0.002 (0.008)	-0.0002 (0.006)	0.570* (0.226)	-0.002 (0.002)	0.013 (0.008)	-0.013+ (0.006)	-0.004 (0.0035)
Hispanic	-0.005 (0.008)	-0.005 (0.004)	-0.001 (0.001)	0.0001 (0.0001)	-0.0005** (0.0002)	-0.002 (0.009)**	-0.0002 (0.0003)	-0.060 (0.161)	-0.0002 (0.003)	0.001 (0.007)	-0.005* (0.002)	0.007** (0.009)**
White	0.014* (0.006)	0.002 (0.002)	0.001 (0.001)	0.0002 (0.0001)	0.00004 (0.0002)	0.009** (0.003)	-0.0003 (0.0004)	0.473*** (0.137)	-0.002 (0.002)	0.003 (0.007)	0.003 (0.003)	0.009*** (0.002)
Other Race	0.034*** (0.009)	0.006* (0.003)	0.002+ (0.001)	0.00001 (0.00004)	0.0004 (0.0002)	0.007* (0.004)	0.0006 (0.0004)	0.243 (0.188)	-0.007* (0.003)	0.004 (0.01)	0.004 (0.004)	-0.0001 (0.005)
<i>No Economic Disadvantage</i>												
Black	0.040*** (0.009)	-0.007 (0.005)	0.001* (0.001)	0.0004* (0.0001)	-0.0001 (0.0002)	0.007+ (0.004)	0.0006* (0.0003)	0.505*** (0.126)	-0.005* (0.002)	0.0001 (0.008)	0.0025 (0.004)	0.006+ (0.003)
Hispanic	0.027*** (0.006)	0.001 (0.001)	0.001 (0.001)	0.0001* (0.0001)	-0.000001 (0.00001)	0.008*** (0.002)	0.0009** (0.0003)	0.573*** (0.120)	0.0008 (0.009)	0.009 (0.003)	0.008* (0.002)	0.009*** (0.006)**
White	0.021*** (0.005)	0.007*** (0.001)	0.002*** (0.001)	0.00002 (0.0001)	-0.00009 (0.0001)	0.007** (0.002)	0.0003 (0.0002)	0.336*** (0.083)	-0.005** (0.002)	0.002 (0.009)	0.005** (0.002)	0.006** (0.002)
Other Race	0.027*** (0.006)	0.011*** (0.002)	0.002*** (0.001)	0.00003 (0.0001)	-0.000001 (0.0002)	0.009** (0.003)	0.0003 (0.0003)	0.308** (0.102)	-0.007* (0.003)	0.014 (0.014)	0.01* (0.005)	0.004 (0.002)
High School - Accept Effects												
<i>Economic Disadvantage</i>												
Black	-0.015** (0.006)	-0.005* (0.002)	-0.003*** (0.001)	-0.0001 (0.0001)	-0.0002 (0.0002)	-0.002 (0.003)	0.0002 (0.0005)	-0.179 (0.259)	0.015*** (0.005)	-0.002 (0.008)	-0.001 (0.002)	-0.011* (0.005)
Hispanic	-0.012+ (0.007)	-0.001 (0.001)	0.0001 (0.001)	-0.0001* (0.00004)	-0.0005** (0.0002)	0.002 (0.002)	-0.0003 (0.0003)	-0.485* (0.207)	0.005 (0.005)	-0.008 (0.005)	-0.002 (0.002)	-0.009** (0.003)
White	0.004 (0.005)	0.001 (0.001)	-0.002 (0.001)	-0.0002+ (0.0001)	-0.0003+ (0.0002)	0.006* (0.002)	-0.001*** (0.0003)	-0.115 (0.149)	0.01** (0.003)	-0.007 (0.006)	-0.001 (0.002)	0.0009 (0.002)
Other Race	0.016 (0.010)	0.003* (0.001)	-0.001 (0.001)	-0.0001 (0.0001)	0.0001 (0.0002)	0.0007 (0.003)	-0.0002 (0.0004)	0.262 (0.244)	-0.0007 (0.003)	-0.003 (0.011)	0.008* (0.004)	0.007+ (0.004)
<i>No Economic Disadvantage</i>												
Black	0.002 (0.008)	-0.001 (0.002)	-0.002*** (0.001)	0.0001 (0.0002)	-0.0001 (0.0001)	-0.0006 (0.003)	-0.0005 (0.0004)	0.118 (0.222)	0.006+ (0.004)	-0.009 (0.01)	0.001 (0.002)	-0.004 (0.003)
Hispanic	0.009 (0.008)	-0.003 (0.002)	-0.004 (0.001)	-0.0001 (0.0001)	-0.0004+ (0.0002)	0.005+ (0.003)	-0.0007** (0.0003)	0.044 (0.191)	0.006 (0.004)	-0.006 (0.007)	0.002 (0.001)	-0.002 (0.003)
White	0.005 (0.004)	0.004*** (0.001)	-0.0003 (0.001)	-0.00002 (0.00004)	-0.0001 (0.0001)	0.003 (0.003)	-0.0004* (0.0002)	-0.094 (0.098)	-0.001 (0.002)	-0.012 (0.008)	0.003** (0.001)	0.002+ (0.001)
Other Race	0.002 (0.007)	0.005*** (0.001)	0.00001 (0.001)	-0.0001 (0.0001)	0.0002 (0.0002)	-0.001 (0.003)	-0.0005 (0.0003)	0.140 (0.156)	-0.007** (0.002)	-0.028* (0.013)	0.002 (0.001)	0.005+ (0.002)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,292,626	11,684,325	11,684,325	11,684,325	11,684,325

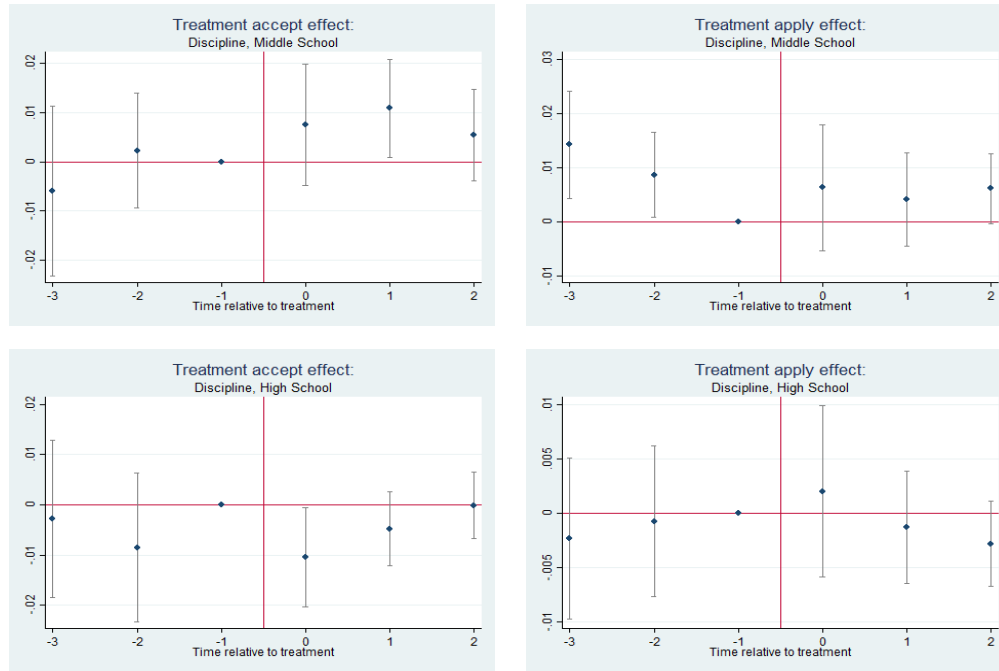
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Table 3.4. While not shown due to space constraints, all models include application variables interacted with the same set of demographic categories as shown above. Demographic interactions in this table are non-additive and represent mutually exclusive student groups.

Table 3.8: Long-term Student Outcomes, by Demographic Group

	(1) Graduate High School	(2) Enroll College	(3) Enroll: 2-year College	(4) Enroll: 4-year College	(5) Graduate College	(6) Graduate: 2-year College	(7) Graduate: 4-year College	(8) Employed	(9) Income	(10) Income Given Employment
Acceptance Exposure Effects										
<i>Economic Disadvantage</i>										
Black	-0.001 (0.007)	-0.025*** (0.006)	-0.043*** (0.006)	0.007 (0.006)	0.008 (0.005)	0.005+ (0.003)	0.003 (0.005)	-0.008 (0.012)	228.2 (427.6)	-313.4 (290.1)
Hispanic	0.002 (0.009)	0.003 (0.008)	0.002 (0.009)	0.007 (0.006)	0.012* (0.006)	0.003 (0.003)	0.01+ (0.005)	0.017* (0.009)	714.1*** (198.2)	637.3* (251.0)
White	-0.05*** (0.011)	-0.028*** (0.011)	-0.027*** (0.01)	-0.01 (0.008)	-0.002 (0.006)	0.004 (0.004)	-0.008 (0.006)	-0.027 (0.018)	-14.4 (484.8)	377.4 (819.6)
Other Race	0.009 (0.018)	0.008 (0.032)	0.018 (0.034)	-0.025 (0.016)	-0.003 (0.031)	0.001 (0.011)	-0.01 (0.025)	0.012 (0.017)	-26.6 (405.1)	-266.3 (554.7)
No Economic Disadvantage										
Black	0.003 (0.01)	-0.035* (0.018)	-0.052** (0.018)	0.002 (0.013)	-0.017** (0.006)	-0.002 (0.003)	-0.016* (0.006)	-0.028** (0.009)	-626.1** (222.5)	-322.2 (280.5)
Hispanic	-0.023 (0.014)	-0.03+ (0.017)	-0.031* (0.012)	-0.003 (0.017)	0.005 (0.012)	0.004 (0.005)	0.001 (0.012)	-0.012 (0.008)	-57.5 (280.3)	476.4 (349.1)
White	-0.02 (0.016)	-0.036 (0.023)	-0.019 (0.02)	-0.032+ (0.017)	-0.023 (0.017)	-0.006 (0.004)	-0.021 (0.017)	-0.015 (0.014)	-406.4 (355.3)	123.0 (237.2)
Other Race	-0.001 (0.016)	0.005 (0.023)	0.041* (0.021)	-0.015 (0.029)	-0.007 (0.031)	-0.01 (0.01)	0.002 (0.032)	-0.006 (0.014)	243.1 (458.4)	936.8 (637.1)
Student Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Table 3.5. While not shown due to space constraints, all models include application variables interacted with the same set of demographic categories as shown above. Demographic interactions in this table are non-additive and represent mutually exclusive student groups.

Figure 3.2: Timing of Grant Acceptance and Application Treatment Effects on Student Discipline



Each of the graphs shows separate coefficient estimates from the same regression. Bars surrounding coefficients represent a 95% confidence interval for each estimate with standard errors clustered at the school district level and the year preceding treatment omitted. Because a school district may receive multiple grants within a sample period, these graphs are created by duplicating the data for each possible treatment year and stacking these data sets to form a "pseudo panel." In each year panel, the designated treatment year is considered over time, and treatments in adjacent years are included as model controls. Due to computational constraints given the size of my data set, these graphs were produced using data collapsed to the school district-grade-year level and weighted by the number of students within these cells. The regressions correspond to the fully specified model, that includes student demographic covariates, student enrollment in a district-grade-year cell, and fixed effects for grade, year, and school district. I include treatments in 2000-2013, so that each centered treatment year has at least one year of observed pre-treatment data.

Table 3.9: Robustness Tests, Short-term Student Discipline Outcomes

	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
1. Baseline Model				
Middle School: Accept	0.010* (0.004)	0.005 (0.004)	0.009* (0.004)	-0.0001 (0.0003)
High School: Accept	-0.007+ (0.004)	-0.010* (0.005)	-0.002 (0.003)	-0.0009*** (0.0003)
Observations	13,596,577	13,596,577	13,596,577	13,596,577
2. Apply Ever				
Middle School: Accept	0.012** (0.004)	0.006 (0.005)	0.010* (0.004)	-0.0001 (0.0003)
High School: Accept	-0.006 (0.004)	-0.009* (0.005)	-0.002 (0.003)	-0.0009*** (0.0003)
Observations	10,698,313	10,698,313	10,698,313	10,698,313
3. Accepted Ever				
Middle School: Accept	0.009+ (0.005)	0.003 (0.005)	0.006 (0.005)	-0.0001 (0.0003)
High School: Accept	-0.004 (0.004)	-0.007 (0.005)	-0.0012 (0.003)	-0.0010*** (0.0003)
Observations	9,052,378	9,052,378	9,052,378	9,052,378
4. Both Accepted and Rejected				
Middle School: Accept	0.012* (0.005)	0.005 (0.005)	0.009+ (0.005)	-0.0001 (0.0003)
High School: Accept	-0.006 (0.005)	-0.009+ (0.005)	-0.004 (0.004)	-0.0010** (0.0003)
Observations	7,419,861	7,419,861	7,419,861	7,419,861
5. Application Group by Year Effects				
Middle School: Accept	0.013** (0.005)	0.007 (0.005)	0.009* (0.004)	-0.0001 (0.0003)
High School: Accept	-0.005 (0.004)	-0.008+ (0.005)	-0.002 (0.003)	-0.0009*** (0.0003)
Observations	13,596,577	13,596,577	13,596,577	13,596,577
6. Unweighted by District Regression				
Middle School: Accept	-0.010 (0.006)	-0.005 (0.007)	-0.012** (0.005)	-0.0015 (0.0015)
High School: Accept	-0.002 (0.005)	0.001 (0.005)	-0.003 (0.003)	-0.0008 (0.0006)
Observations	62,990	62,990	62,990	62,990

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Table 3.3. Specifications 2-4 show regression results restricted to school districts that applied for a grant during the sample period, were accepted for a grant during the period, or were both accepted and rejected for grants during the sample period. Specification 5 replicates the baseline model using data collapsed at the district-grade-year level that is not weighted by the number of students in a school district-grade-year.

Table 3.10: Robustness Tests, Long-term student outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Graduate High School	Enroll College	Enroll: 2-year College	Enroll: 4-year College	Graduate College	Graduate: 2-year College	Graduate: 4-year College	Employed	Income	Income Given Employment
1. Baseline Model										
Accept Exposure	-0.017* (0.007)	-0.029** (0.011)	-0.037** (0.011)	-0.001 (0.005)	-0.0001 (0.004)	-0.0007 (0.002)	-0.002 (0.005)	0.001 (0.006)	355.1 (295.0)	558.4 (406.1)
Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455
2. Apply Ever										
Accept Exposure	-0.014+ (0.007)	-0.025* (0.011)	-0.032** (0.011)	0.001 (0.006)	0.0003 (0.004)	-0.0005 (0.002)	-0.001 (0.004)	0.002 (0.006)	324.3 (293.8)	513.9 (412.7)
Observations	1,980,714	1,980,714	1,980,714	1,980,714	710,375	710,375	710,375	710,375	710,375	424,374
3. Accept Ever										
Accept Exposure	-0.009 (0.009)	-0.022+ (0.013)	-0.027* (0.013)	0.001 (0.008)	0.0003 (0.004)	-0.0007 (0.002)	-0.001 (0.004)	0.002 (0.006)	310.2 (298.9)	506.6 (421.0)
Observations	1,675,383	1,675,383	1,675,383	1,675,383	598,308	598,308	598,308	598,308	598,308	353,618
4. Both Accepted and Rejected										
Accept Exposure	-0.007 (0.011)	-0.022 (0.015)	-0.030+ (0.015)	0.006 (0.008)	0.003 (0.004)	-0.0011 (0.002)	0.003 (0.004)	-0.002 (0.008)	254.9 (432.5)	520.6 (623.7)
Observations	1,374,692	1,374,692	1,374,692	1,374,692	489,384	489,384	489,384	489,384	489,384	285,445
5. Application Group by Year Effects										
Accept Exposure	-0.007 (0.009)	-0.022+ (0.013)	-0.026* (0.013)	0.001 (0.008)	0.0007 (0.004)	0.0003 (0.002)	-0.002 (0.005)	0.001 (0.006)	392.1 (308.1)	648.0 (426.7)
Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455
6. Unweighted by District Regression										
Accept Exposure	-0.002 (0.011)	-0.013 (0.011)	-0.029* (0.012)	0.008 (0.009)	-0.002 (0.013)	-0.007 (0.007)	0.001 (0.010)	-0.006 (0.013)	-376.1 (597.5)	-410.6 (847.2)
Observations	8,861	8,861	8,861	8,861	3,217	3,217	3,217	3,217	3,217	3,192

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Table 3.5. Specifications 2-4 show regression results restricted to school districts that applied for a grant during the sample period, were accepted for a grant during the period, or were both accepted and rejected for grants during the sample period. Specification 5 replicates the baseline model using data collapsed at the district-cohort-year level that is not weighted by the number of students in a school district-cohort-year.

Appendices

Appendix A

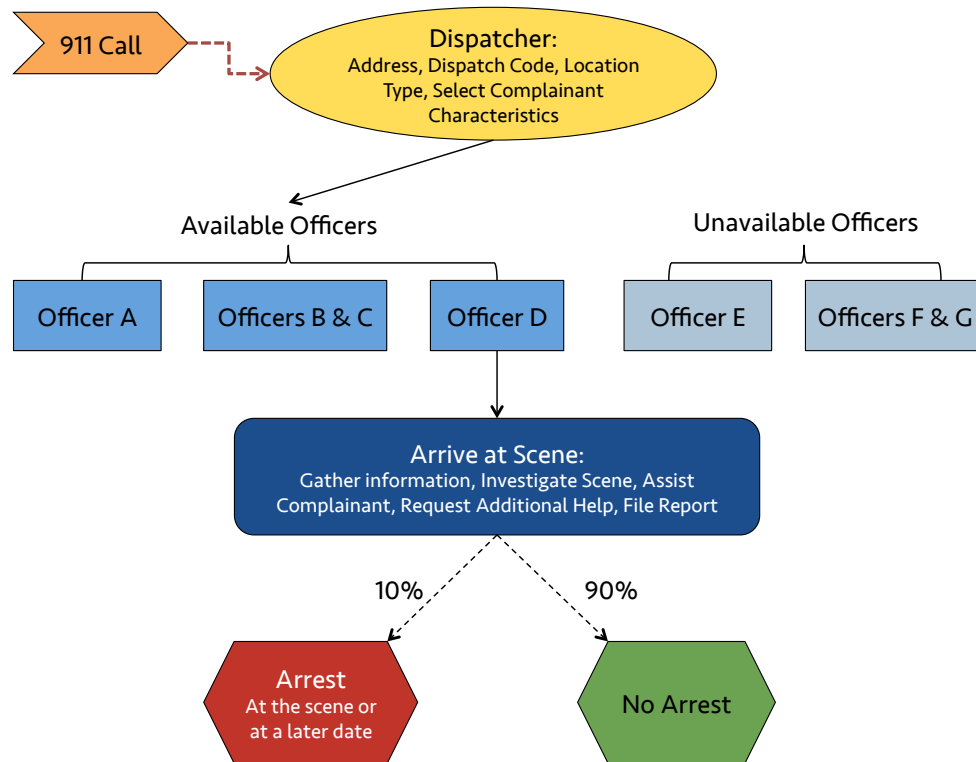
Appendix:

“Whose Help is on the Way?”

The Importance of Individual Police Officers in
Law Enforcement Outcomes

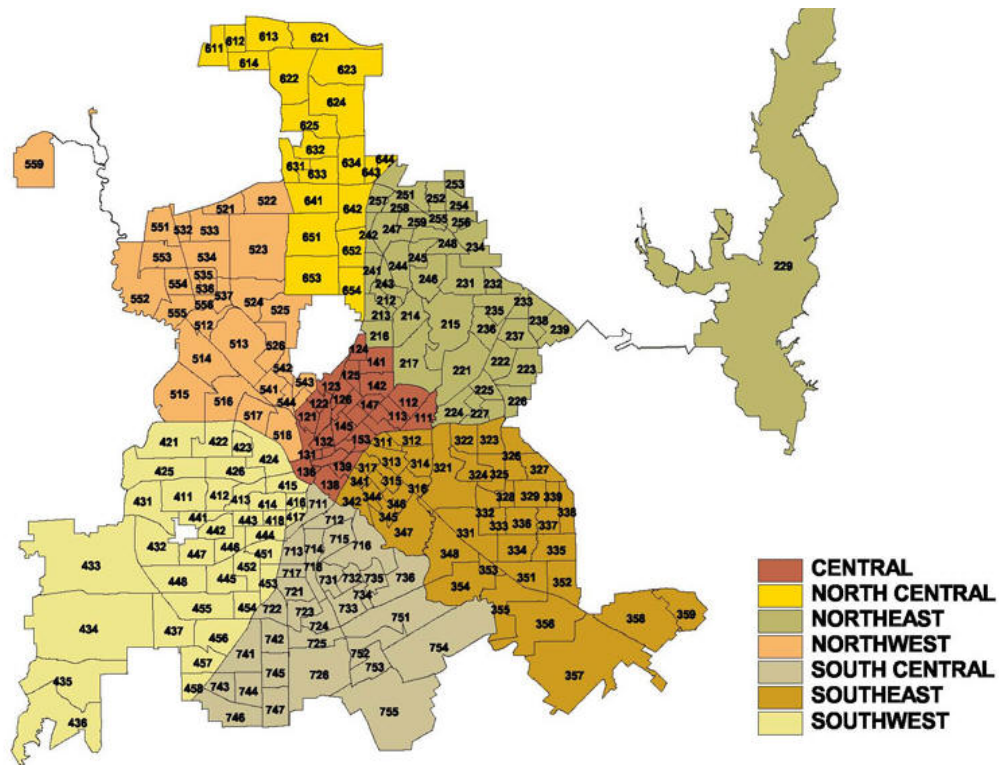
A.1 Appendix Tables and Figures

Figure A.1: Steps involved in an Incident Response



This figure displays an outline of an incident response path at the Dallas Police Department. Information on call response protocols was obtained through conversations with officers and dispatchers at the department.

Figure A.2: Police Beats and Police Divisions in Dallas, TX



This figure shows a map of the 234 police beats contained in the 7 police divisions in Dallas. Police sectors are geographic units that are collections of beats within police divisions (35 total sectors). Map was obtained from the North Dallas Neighborhood Alliance: <http://www.ndna-tx.org/crimeWatch/dallasPolice/DivMap.aspx>.

Table A.1: Summary Statistics, Officer Sorting Robustness Samples

Table A.1.A: Summary Statistics: Outcomes, Officers, and Complainants

	<i>Analysis Sample</i>		<i>Low Availability Sample</i>		<i>Unlikely Response Sample</i>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Total Observations	217,633		116,070		102,203	
Total Incidents	161,531		86,091		87,139	
Total Officers	1,851		1,556		1,505	
Outcomes						
Arrest	0.11	(0.31)	0.11	(0.31)	0.09	(0.29)
Suspect	0.19	(0.39)	0.19	(0.39)	0.17	(0.37)
Felony Arrest	0.04	(0.18)	0.03	(0.18)	0.03	(0.17)
Misdemeanor Arrest	0.08	(0.27)	0.07	(0.26)	0.07	(0.25)
Arrestee Black	0.04	(0.19)	0.04	(0.19)	0.03	(0.18)
Arrestee Hispanic	0.01	(0.12)	0.01	(0.12)	0.01	(0.12)
Arrestee White	0.01	(0.12)	0.01	(0.12)	0.01	(0.11)
Officer Characteristics						
Officer Arrest Rate	0.11	(0.05)	0.11	(0.05)	0.11	(0.05)
Two Responders	0.52	(0.50)	0.51	(0.50)	0.31	(0.46)
Black	0.24	(0.39)	0.24	(0.40)	0.24	(0.41)
Hispanic	0.21	(0.35)	0.21	(0.35)	0.21	(0.37)
White	0.50	(0.45)	0.50	(0.45)	0.50	(0.47)
Female	0.16	(0.32)	0.16	(0.32)	0.16	(0.34)
Age	38.37	(9.30)	38.30	(9.35)	38.53	(9.75)
Trainee	0.07	(0.20)	0.07	(0.20)	0.06	(0.20)
Sergeant	0.01	(0.10)	0.01	(0.08)	0.01	(0.08)
Salary (\$10,000s)	5.88	(1.02)	5.87	(1.02)	5.89	(1.08)
Years of Experience	11.82	(8.56)	11.73	(8.60)	11.92	(8.98)
Total Incidents	171.03	(103.62)	177.34	(101.79)	179.26	(101.22)
Complainant Characteristics						
Victim with Injury	0.10	(0.29)	0.09	(0.29)	0.08	(0.28)
Number of Complainants	1.59	(0.99)	1.59	(0.99)	1.56	(0.97)
Black	0.36	(0.48)	0.36	(0.48)	0.35	(0.48)
Hispanic	0.35	(0.48)	0.35	(0.48)	0.35	(0.48)
White	0.32	(0.47)	0.32	(0.47)	0.32	(0.47)
Female	0.50	(0.50)	0.50	(0.50)	0.50	(0.50)

These tables display summary statistics for covariates used in analysis. The first column, “Analysis Sample”, summarizes the primary analysis sample. The second column, “Low Availability Sample,” consists of observations where fewer officers are available to respond to at the time of the incident. The third column, “Unlikely Response Sample,” consists of observations where the predicted probability that an officer responds to the incident is low, given other observables in the model. More details on the construction of these samples can be found in Section 1.5.1.1. Officer arrest rate and number of incidents is calculated over all observations in the raw data. Complainant characteristics are shown for observations that have information on complainant demographics, which corresponds to 95% of the sample. Information for suspect outcomes is only available prior to 2017 and is summarized for the period available. Robustness samples are restricted to officers with at least 25 observations within the relevant sub-sample.

Table A.1.B: Summary Statistics: Incident Urgency, Location Type, and Dispatch Code

	<i>Analysis Sample</i>		<i>Low Availability Sample</i>		<i>Unlikely Response Sample</i>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Total Observations	217,633		116,070		102,203	
Total Incidents	161,531		86,091		87,139	
Total Officers	1,851		1,556		1,505	
<i>Call Urgency</i>						
Time to Dispatch (Minutes)	24.09	(27.94)	24.20	(27.96)	25.31	(28.54)
<i>Location Type</i>						
Apartment	0.13	(0.33)	0.13	(0.33)	0.12	(0.33)
Residence Other	0.15	(0.36)	0.15	(0.36)	0.15	(0.36)
Bar/Club/Entertainment	0.03	(0.17)	0.03	(0.17)	0.03	(0.17)
Retail	0.07	(0.25)	0.07	(0.25)	0.06	(0.24)
Business Other	0.05	(0.22)	0.05	(0.22)	0.05	(0.21)
Govt/Health/School/Religion	0.01	(0.10)	0.01	(0.10)	0.01	(0.10)
Motor Vehicle	0.02	(0.15)	0.02	(0.14)	0.02	(0.14)
Parking Lot	0.24	(0.43)	0.24	(0.43)	0.25	(0.43)
Street	0.20	(0.40)	0.20	(0.40)	0.21	(0.40)
Outdoor Other	0.05	(0.22)	0.05	(0.22)	0.05	(0.22)
Other Location	0.05	(0.21)	0.05	(0.21)	0.05	(0.21)
<i>Dispatch Code Type</i>						
Criminal Assault	0.01	(0.12)	0.02	(0.12)	0.01	(0.11)
Armed Encounter/Active Shooter	0.02	(0.13)	0.02	(0.13)	0.02	(0.12)
Injured Person	0.01	(0.11)	0.01	(0.11)	0.01	(0.11)
Robbery	0.06	(0.24)	0.06	(0.24)	0.05	(0.22)
Burglary of Business	0.05	(0.22)	0.05	(0.22)	0.05	(0.22)
Burglary of Residence	0.11	(0.32)	0.11	(0.32)	0.12	(0.32)
Burglary of Motor Vehicle	0.17	(0.38)	0.17	(0.38)	0.18	(0.39)
Unauthorized Use of Motor Vehicle	0.05	(0.22)	0.05	(0.23)	0.06	(0.23)
Theft	0.07	(0.26)	0.07	(0.26)	0.08	(0.27)
Criminal Mischief	0.07	(0.25)	0.07	(0.26)	0.07	(0.26)
Major Disturbance	0.11	(0.32)	0.11	(0.32)	0.10	(0.30)
Major Accident	0.04	(0.21)	0.04	(0.20)	0.04	(0.20)
Minor Accident	0.09	(0.29)	0.09	(0.29)	0.10	(0.30)
Other - Serious	0.07	(0.26)	0.07	(0.26)	0.06	(0.24)
Other - Minor	0.04	(0.21)	0.04	(0.21)	0.04	(0.20)

These tables display summary statistics for covariates used in analysis. The first column, “Analysis Sample”, summarizes the primary analysis sample. The second column, “Low Availability Sample,” consists of observations where fewer officers are available to respond to at the time of the incident. The third column, “Unlikely Response Sample,” consists of observations where the predicted probability that an officer responds to the incident is low, given other observables in the model. More details on the construction of these samples can be found in Section 1.5.1.1. Call Urgency or Time to Dispatch (Minutes) is a variable that captures the priority level or severity of a given call at the time of dispatch. Robustness samples are restricted to officers with at least 25 observations within the relevant sub-sample.

Table A.2: Additional Policing Outcomes

Table A.2.A: Summary Results, Additional Policing Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Arrest	Suspect	Felony Arrest	Misdemeanor Arrest	Arrest, Single Responder	Arrest, Urgent Calls	Arrest, Non-Urgent Calls	Time to Clearance (Minutes)
% of R-2 from Officer Effects	16.5%	12.2%	26.0%	15.6%	11.9%	21.5%	14.1%	32.4%
% of Adj. R-2 from Officer Effects	11.9%	7.3%	20.1%	10.3%	9.7%	16.1%	7.2%	9.1%
S.D. of Officer Effect	0.036	0.046	0.020	0.029	0.031	0.053	0.035	46.5
% Change: 1 S.D. Increase in Officer Effect	32.8%	24.2%	55.7%	37.6%	51.8%	39.2%	44.0%	31.6%
Gap: 10th to 90th Percentile in Officer Effect	0.082	0.111	0.046	0.068	0.071	0.126	0.085	112.0
% Change: 10th to 90th Percentile in Officer Effect	75.9%	58.3%	131.2%	89.2%	119.1%	93.7%	106.1%	76.1%
Correlation to Arrest Officer Effect		0.688	0.555	0.806	0.719	0.844	0.606	0.057
Mean of Outcome	0.109	0.191	0.035	0.076	0.060	0.134	0.080	147.2
S.D. of Outcome	0.311	0.393	0.185	0.266	0.237	0.341	0.271	765.4
Total Officers	1,851	1,739	1,851	1,851	1,184	1,546	1,444	1,851
Total Observations	217,633	160,872	217,633	217,633	96,696	105,881	99,166	217,549

This table summarizes the analysis results across additional policing outcomes using the full sample. Column (1) replicates the results for the main arrest outcome, column (2) measures whether a suspect was identified for the incident, columns (3) and (4) measure whether the incident resulted in a felony or misdemeanor arrest, and column (5) includes the sub-sample of incident observations with only one officer responder. Columns (6) and (7), measure arrest outcomes for urgent and not urgent calls, where these samples are defined by whether a call is above or below the median amount of time between when a call is placed and an officer is dispatched (call severity variable). Lastly, column (8) measures the outcome of time an officer spends responding to an incident measured by the minutes passed between dispatch and the clearance of a call (or time when the initial report is filed by the officer). Each specification includes officers with at least 25 observations for the relevant sample.

Table A.2.B: Correlation of Officer Fixed Effects, Additional Policing Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Arrest	Suspect	Felony Arrest	Misdemeanor Arrest	Arrest, Single Responder	Arrest, Urgent Calls	Arrest, Non-Urgent Calls	Time to Clearance (Minutes)
(1) Arrest	1.00							
(2) Suspect	0.69	1.00						
(3) Felony Arrest	0.55	0.41	1.00					
(4) Misdemeanor Arrest	0.81	0.53	0.06	1.00				
(5) Arrest, Single Responder	0.72	0.57	0.42	0.55	1.00			
(6) Arrest, Urgent Calls	0.84	0.58	0.53	0.66	0.61	1.00		
(7) Arrest, Non-Urgent Calls	0.61	0.39	0.31	0.55	0.50	0.14	1.00	
(8) Time to Clearance (Minutes)	0.06	0.03	0.08	-0.01	0.09	0.05	0.02	1.00

This table summarizes the analysis results across additional policing outcomes using the full sample. Column (1) replicates the results for the main arrest outcome, column (2) measures whether a suspect was identified for the incident, columns (3) and (4) measure whether the incident resulted in a felony or misdemeanor arrest, and column (5) includes the sub-sample of incident observations with only one officer responder. Columns (6) and (7), measure arrest outcomes for urgent and not urgent calls, where these samples are defined by whether a call is above or below the median amount of time between when a call is placed and an officer is dispatched (call severity variable). Lastly, column (8) measures the outcome of time an officer spends responding to an incident measured by the minutes passed between dispatch and the clearance of a call (or time when the initial report is filed by the officer). Panel (B) shows the correlations between officer fixed effects obtained from the additional policing outcomes in Panel (A). Correlations shown in the bottom panel are calculated for overlapping observations across different specifications. Information for suspect outcomes is only available prior to 2017. Each specification includes officers with at least 25 observations for the relevant sample.

Table A.3: Officer Effects and Officer Demographics, Additional Policing Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Outcome: Officer Effect</i>	Arrest	Suspect	Felony Arrest	Misdemeanor Arrest	Arrest, Single Responder	Arrest, Urgent Calls	Arrest, Non-Urgent Calls	Time to Clearance (Minutes)
Black	0.002 (0.004)	-0.004 (0.005)	0.0004 (0.002)	0.002 (0.003)	0.0001 (0.004)	0.001 (0.007)	0.001 (0.005)	8.21 (5.39597)
Hispanic	-0.002 (0.004)	-0.008 (0.006)	0.001 (0.002)	-0.002 (0.003)	-0.003 (0.005)	0.002 (0.007)	-0.006 (0.005)	4.70 (5.45542)
White	0.004 (0.004)	0.003 (0.005)	0.003 (0.002)	0.000 (0.003)	0.003 (0.004)	0.006 (0.006)	0.001 (0.005)	0.661 (5.15985)
Female	-0.004+ (0.002)	-0.002 (0.003)	-0.003+ (0.001)	-0.0009 (0.002)	-0.004* (0.002)	-0.002 (0.003)	-0.001 (0.002)	-0.383 (2.86040)
Age	-0.0004* (0.0002)	-0.0003 (0.0002)	-0.0002* (0.0001)	-0.0001 (0.0001)	-0.0002 (0.0002)	-0.0001 (0.0003)	-0.0004* (0.0002)	0.711** (0.22680)
Trainee	-0.010** (0.003)	-0.010* (0.005)	-0.006*** (0.002)	-0.005* (0.003)	-0.005 (0.006)	-0.003 (0.005)	-0.011** (0.004)	2.01 (4.00223)
Sergeant	-0.005 (0.005)	-0.014+ (0.007)	-0.003 (0.003)	-0.002 (0.004)	-0.009+ (0.005)	-0.001 (0.008)	-0.017* (0.008)	-7.23 (6.14398)
Experience	0.0019*** (0.0004)	0.002*** (0.0006)	0.0004+ (0.0002)	0.001** (0.0004)	0.0015** (0.0005)	0.0022** (0.0007)	0.0018*** (0.0004)	-0.760 (0.600)
Experience^2	-0.00004*** (0.00001)	-0.00005*** (0.00001)	-0.00001 (0.00001)	-0.00003** (0.00001)	-0.00003** (0.00001)	-0.00005** (0.00002)	-0.00004*** (0.00001)	-0.002 (0.015)
Observations	1,832	1,736	1,832	1,832	1,181	1,541	1,441	1,832
R-squared	0.041	0.029	0.021	0.021	0.023	0.021	0.038	0.013
Fixed Effect Mean	-0.001	-0.001	-0.002	-0.002	-0.003	-0.002	-0.002	-8.564
Fixed Effect S.D.	0.036	0.046	0.020	0.029	0.031	0.053	0.035	46.350
Outcome Mean	0.109	0.191	0.035	0.076	0.060	0.134	0.080	147.200
Outcome S.D.	0.311	0.393	0.185	0.266	0.237	0.341	0.271	765.400

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table shows regression results of officer effects regressed on fixed officer characteristics, at the officer level. Each column uses the officer effects derived from the specifications in Table A.2. Robust standard errors are in parentheses. Other race officers are the omitted race category. Officers without demographic information are excluded from the regressions. Information for suspect outcomes is only available prior to 2017. Each specification includes officers with at least 25 observations for the relevant sample.

Table A.4: Officer Effects and Officer Demographics, Robustness Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Outcome: Officer Effect</i>	Base Model	Add Reporting Area Fixed Effects	Add Sector*Month Fixed Effects	Add Individual Shift Effects	Add Full Set of Dispatch Codes	First Stage Fixed Effects: >100 N	First Stage Fixed Effects: Weighted by N	First Stage Fixed Effects: Unadjusted
Black	0.002 (0.004)	0.001 (0.004)	0.001 (0.004)	0.002 (0.005)	0.002 (0.004)	0.005 (0.006)	0.0002 (0.005)	0.003 (0.005)
Hispanic	-0.002 (0.004)	-0.003 (0.004)	-0.002 (0.004)	-0.001 (0.005)	-0.001 (0.004)	0.001 (0.006)	-0.003 (0.005)	-0.002 (0.005)
White	0.004 (0.004)	0.003 (0.004)	0.004 (0.004)	0.005 (0.005)	0.004 (0.004)	0.005 (0.006)	0.002 (0.004)	0.006 (0.005)
Female	-0.004+ (0.002)	-0.004+ (0.002)	-0.004+ (0.002)	-0.004+ (0.003)	-0.004* (0.002)	-0.002 (0.003)	-0.002 (0.002)	-0.005+ (0.003)
Age	-0.0004* (0.0002)	-0.0004* (0.0002)	-0.0003* (0.0002)	-0.0005* (0.0002)	-0.0002 (0.0002)	-0.0004 (0.0002)	-0.0003 (0.0002)	-0.0006** (0.0002)
Trainee	-0.010** (0.003)	-0.010** (0.003)	-0.010** (0.003)	-0.012** (0.004)	-0.008** (0.003)	-0.007 (0.005)	-0.010** (0.004)	-0.014** (0.004)
Sergeant	-0.005 (0.005)	-0.008 (0.005)	-0.004 (0.005)	-0.003 (0.006)	-0.006 (0.005)	-0.010 (0.007)	-0.005 (0.006)	-0.005 (0.007)
Experience	0.0019*** (0.0004)	0.0021*** (0.0004)	0.0019*** (0.0004)	0.0028*** (0.0005)	0.0019*** (0.0004)	0.0027*** (0.0006)	0.0022*** (0.0005)	0.0029*** (0.0006)
Experience^2	-0.00004*** (0.00001)	-0.00004*** (0.00001)	-0.00004*** (0.00001)	-0.00006*** (0.00001)	-0.00004*** (0.00001)	-0.00005*** (0.00002)	-0.00005*** (0.00001)	-0.00006*** (0.00001)
Observations	1,832	1,832	1,832	1,832	1,832	941	1,832	1,832
R-squared	0.041	0.042	0.039	0.047	0.043	0.043	0.030	0.040
Fixed Effect Mean	-0.001	-0.001	-0.001	0.000	-0.001	0.000	0.000	0.001
Fixed Effect S.D.	0.036	0.036	0.036	0.044	0.033	0.039	0.041	0.050
Outcome Mean	0.109	0.109	0.109	0.109	0.109	0.103	0.109	0.109
Outcome S.D.	0.311	0.311	0.311	0.311	0.311	0.305	0.311	0.311

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table shows regression results of officer effects regressed on fixed officer characteristics, at the officer level. Each column uses the officer effects derived from the specifications in Table A.2. Robust standard errors are in parentheses. Other race officers are the omitted race category. Officers without demographic information are excluded from the regressions.

Table A.5: Summary Statistics, Racial Bias Test Sample

Table A.5.A: Summary Statistics: Outcomes, Officers, and Complainants

	<i>Racial Bias Test Sample</i>		<i>Black Officers</i>		<i>Hispanic Officers</i>		<i>White Officers</i>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Total Observations	156,361		39,770		26,571		85,573	
Total Incidents	129,482		34,020		22,560		68,551	
Total Officers	1,613		423		314		828	
Outcomes								
Arrest	0.09	(0.29)	0.09	(0.28)	0.08	(0.27)	0.10	(0.30)
Arrestee Black	0.03	(0.18)	0.04	(0.20)	0.02	(0.15)	0.03	(0.18)
Arrestee Hispanic	0.01	(0.11)	0.01	(0.10)	0.01	(0.11)	0.01	(0.12)
Arrestee White	0.01	(0.11)	0.01	(0.09)	0.01	(0.09)	0.02	(0.12)
Officer Characteristics								
Officer Arrest Rate	0.10	(0.05)	0.10	(0.05)	0.10	(0.05)	0.11	(0.06)
Two Responders	0.34	(0.47)	0.29	(0.45)	0.30	(0.46)	0.40	(0.49)
Black	0.25	(0.44)	1.00					
Hispanic	0.17	(0.38)			1.00			
White	0.55	(0.50)					1.00	
Female	0.16	(0.34)	0.25	(0.40)	0.15	(0.32)	0.12	(0.30)
Age	39.45	(9.75)	40.69	(9.91)	37.00	(9.63)	39.73	(9.68)
Trainee	0.04	(0.18)	0.04	(0.17)	0.05	(0.19)	0.05	(0.17)
Sergeant	0.01	(0.10)	0.02	(0.13)	0.01	(0.08)	0.01	(0.08)
Salary (\$10,000s)	6.00	(1.07)	6.02	(1.09)	5.80	(1.04)	6.06	(1.08)
Years of Experience	12.87	(9.08)	13.01	(8.70)	10.79	(8.71)	13.56	(9.40)
Total Incidents	181.02	(108.91)	166.86	(103.42)	175.25	(97.01)	187.88	(114.99)
Complainant Characteristics								
Victim with Injury	0.08	(0.27)	0.07	(0.25)	0.08	(0.27)	0.09	(0.29)
Number of Complainants	1.55	(0.94)	1.50	(0.85)	1.50	(0.92)	1.59	(0.99)
Black	0.35	(0.48)	0.46	(0.50)	0.31	(0.46)	0.31	(0.46)
Hispanic	0.35	(0.48)	0.30	(0.46)	0.43	(0.49)	0.34	(0.47)
White	0.33	(0.47)	0.25	(0.43)	0.29	(0.45)	0.37	(0.48)
Female	0.49	(0.50)	0.51	(0.50)	0.49	(0.50)	0.48	(0.50)

These tables display summary statistics the racial bias test analysis. The first column, “Racial Bias Test Sample”, summarizes the primary analysis sample. This sample is restricted to observations where there is a single race for responding officers (if there are multiple officers) and each officer has more than 25 observations. Officer arrest rate and number of incidents is calculated over all observations in the raw data. Columns (2) - (4) describe characteristics of observations for Black, Hispanic, and White officers within the “Racial Bias Test Sample.” Complainant characteristics are shown for observations that have information on complainant demographics, which corresponds to 95% of the sample.

Table A.5.B: Summary Statistics: Incident Urgency, Location Type, and Dispatch Code

	<i>Racial Bias Test Sample</i>		<i>Black Officers</i>		<i>Hispanic Officers</i>		<i>White Officers</i>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Total Observations	156,361		39,770		26,571		85,573	
Total Incidents	129,482		34,020		22,560		68,551	
Total Officers	1,613		423		314		828	
<i>Call Urgency</i>								
Time to Dispatch (Minutes)	24.99	(28.37)	25.56	(28.66)	26.17	(29.02)	24.16	(27.89)
<i>Location Type</i>								
Apartment	0.12	(0.33)	0.13	(0.34)	0.12	(0.32)	0.12	(0.32)
Residence Other	0.15	(0.36)	0.17	(0.37)	0.17	(0.38)	0.14	(0.35)
Bar/Club/Entertainment	0.03	(0.17)	0.03	(0.16)	0.03	(0.16)	0.03	(0.17)
Retail	0.06	(0.24)	0.06	(0.24)	0.05	(0.23)	0.07	(0.25)
Business Other	0.05	(0.22)	0.04	(0.19)	0.05	(0.21)	0.06	(0.23)
Govt/Health/School/Religion	0.01	(0.10)	0.01	(0.10)	0.01	(0.11)	0.01	(0.10)
Motor Vehicle	0.02	(0.14)	0.03	(0.17)	0.02	(0.13)	0.02	(0.13)
Parking Lot	0.25	(0.44)	0.24	(0.43)	0.25	(0.43)	0.26	(0.44)
Street	0.20	(0.40)	0.16	(0.37)	0.20	(0.40)	0.22	(0.42)
Outdoor Other	0.05	(0.22)	0.07	(0.26)	0.07	(0.25)	0.04	(0.19)
Other Location	0.05	(0.21)	0.06	(0.24)	0.04	(0.20)	0.04	(0.20)
<i>Dispatch Code Type</i>								
Criminal Assault	0.01	(0.11)	0.01	(0.10)	0.01	(0.11)	0.01	(0.12)
Armed Encounter/Active Shooter	0.01	(0.12)	0.01	(0.12)	0.02	(0.12)	0.01	(0.12)
Injured Person	0.01	(0.11)	0.01	(0.12)	0.01	(0.10)	0.01	(0.11)
Robbery	0.05	(0.22)	0.05	(0.21)	0.05	(0.22)	0.05	(0.23)
Burglary of Business	0.05	(0.22)	0.05	(0.21)	0.05	(0.23)	0.06	(0.23)
Burglary of Residence	0.12	(0.32)	0.13	(0.33)	0.13	(0.33)	0.11	(0.31)
Burglary of Motor Vehicle	0.19	(0.39)	0.19	(0.39)	0.19	(0.40)	0.18	(0.39)
Unauthorized Use of Motor Vehicle	0.06	(0.23)	0.06	(0.24)	0.06	(0.24)	0.05	(0.22)
Theft	0.08	(0.27)	0.10	(0.30)	0.08	(0.27)	0.07	(0.26)
Criminal Mischief	0.07	(0.26)	0.09	(0.29)	0.08	(0.27)	0.07	(0.25)
Major Disturbance	0.10	(0.29)	0.11	(0.31)	0.09	(0.29)	0.09	(0.29)
Major Accident	0.04	(0.20)	0.03	(0.17)	0.04	(0.21)	0.05	(0.22)
Minor Accident	0.10	(0.30)	0.07	(0.26)	0.09	(0.28)	0.11	(0.32)
Other - Serious	0.06	(0.24)	0.05	(0.22)	0.05	(0.22)	0.07	(0.25)
Other - Minor	0.04	(0.20)	0.05	(0.21)	0.04	(0.20)	0.04	(0.20)

These tables display summary statistics the racial bias test analysis. The first column, “Racial Bias Test Sample”, summarizes the primary analysis sample. This sample is restricted to observations where there is a single race for responding officers (if there are multiple officers) and each officer has more than 25 observations. Columns (2) - (4) describe characteristics of observations for Black, Hispanic, and White officers within the “Racial Bias Test Sample.” Complainant characteristics are shown for observations that have information on complainant demographics, which corresponds to 95% of the sample. Call Urgency or Time to Dispatch (Minutes) is a variable that captures the priority level or severity of a given call at the time of dispatch.

Table A.6: Racial Bias Test, Officer Characteristics in First Stage

Table A.6.A: Racial Bias Test, Officer Characteristics in First Stage, Full Sample

		(1)	(2)	(3)	(4)
<i>Outcome=Officer Effects</i>		Arrest	Arrest Black	Arrest Hispanic	Arrest White
A. Full Sample					
Black Officer		0.0004 (0.005)	0.002 (0.003)	0.002 (0.002)	0.001 (0.002)
Hispanic Officer		-0.004 (0.005)	-0.001 (0.002)	-0.0002 (0.002)	0.001 (0.002)
White Officer		0.003 (0.005)	0.002 (0.002)	0.002 (0.002)	0.001 (0.001)
Black=Hispanic:	F-Test	1.61	2.72	4.93	0.01
	P-Value	0.20	0.10	0.03	0.92
Black=White:	F-Test	0.89	0.24	0.37	0.23
	P-Value	0.35	0.63	0.55	0.63
Hispanic=White:	F-Test	5.10	6.63	3.52	0.11
	P-Value	0.02	0.01	0.06	0.74
Observations		156,361	156,361	156,361	156,361
Arrest Mean		0.091	0.033	0.013	0.012

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table is comparable to Table 1.5. Here, the regressions omit officer fixed effects, θ_i , and measure the direct effect of officer demographic characteristics in the first stage regression. Each regression controls for demographic characteristics of i and includes fixed effects for co-responders, θ_j . Standard errors are clustered at the level of the focal officer, i , and the shift cell, δ_{gt} , to account for error correlations within officers and shifts. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Standard errors are clustered at the focal officer level, θ_i , and the shift cell, δ_{gt} . The full sample, Panel (A), is restricted to observations where responding officers have a single race, and each officer has more than 25 observations within this restriction. The interaction of arrestee outcome race and officer race through the regression coefficients represents a test of officer and arrestee race interaction effects. F-Tests measure whether officers of different races are more likely to make arrests of individuals of different races.

Table A.6.B: Racial Bias Test, Officer Characteristics in First Stage, Full Sample

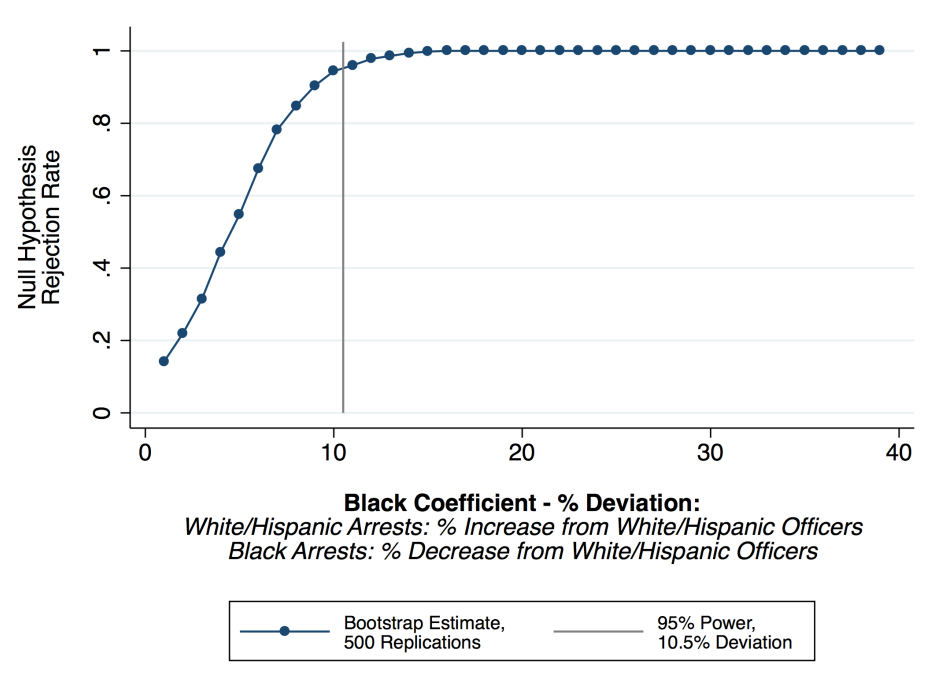
Outcome=Officer Effects		(1)	(2)	(3)	(4)	Outcome=Officer Effects		(5)	(6)	(7)	(8)
B. "Low Availability" Sample		Arrest	Arrest Black	Arrest Hispanic	Arrest White	C. "Unlikely Response" Sample		Arrest	Arrest Black	Arrest Hispanic	Arrest White
Black Officer	Outcome=Officer Effects	-0.002 (0.006)	-0.001 (0.004)	0.002 (0.002)	0.003 (0.002)	Black Officer	Outcome=Officer Effects	0.003 (0.006)	0.003 (0.003)	0.002 (0.002)	0.002 (0.002)
Hispanic Officer	Outcome=Officer Effects	-0.007 (0.007)	-0.007+ (0.004)	0.0003 (0.002)	0.003 (0.002)	Hispanic Officer	Outcome=Officer Effects	-0.002 (0.006)	-0.002 (0.003)	0.000 (0.002)	0.002 (0.002)
White Officer	Outcome=Officer Effects	0.001 (0.006)	-0.001 (0.004)	0.002 (0.002)	0.003 (0.002)	White Officer	Outcome=Officer Effects	0.003 (0.005)	0.002 (0.003)	0.001 (0.002)	0.002 (0.002)
Black=Hispanic:	F-Test	1.49	6.57	2.00	0.14	Black=Hispanic:	F-Test	1.24	7.91	1.86	0.00
	P-Value	0.22	0.01	0.16	0.71		P-Value	0.27	0.00	0.17	0.97
Black=White:	F-Test	0.87	0.01	0.00	0.01	Black=White:	F-Test	0.08	0.25	0.15	0.27
	P-Value	0.35	0.93	0.97	0.93		P-Value	0.78	0.62	0.70	0.60
Hispanic=White:	F-Test	4.47	9.13	2.39	0.23	Hispanic=White:	F-Test	2.37	9.66	1.47	0.26
	P-Value	0.03	0.00	0.12	0.63		P-Value	0.12	0.00	0.23	0.61
Observations		80,874	80,874	80,874	80,874	Observations		81,399	81,399	81,399	81,399
Arrest Mean		0.089	0.032	0.013	0.013	Arrest Mean		0.077	0.028	0.011	0.010

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

This table is comparable to Table 1.5. Here, the regressions omit officer fixed effects, θ_i , and measure the direct effect of officer demographic characteristics in the first stage regression. Each regression controls for demographic characteristics of i and includes fixed effects for co-responders, θ_j . Standard errors are clustered at the level of the focal officer, i , and the shift cell, δ_{gt} , to account for error correlations within officers and shifts. Each arrestee race outcome is defined unconditionally as 1 if an individual of that race was arrested and 0 otherwise. Standard errors are clustered at the focal officer level, θ_i , and the shift cell, δ_{gt} . Panels (B) and (C) represent the overlap of Panel (A) with the "Low Availability" and "Unlikely Response" samples, where the sample is restricted to officers with at least 25 observations within each subset. The interaction of arrestee outcome race and officer race through the regression coefficients represents a test of officer and arrestee race interaction effects. F-Tests measure whether officers of different races are more likely to make arrests of individuals of different races.

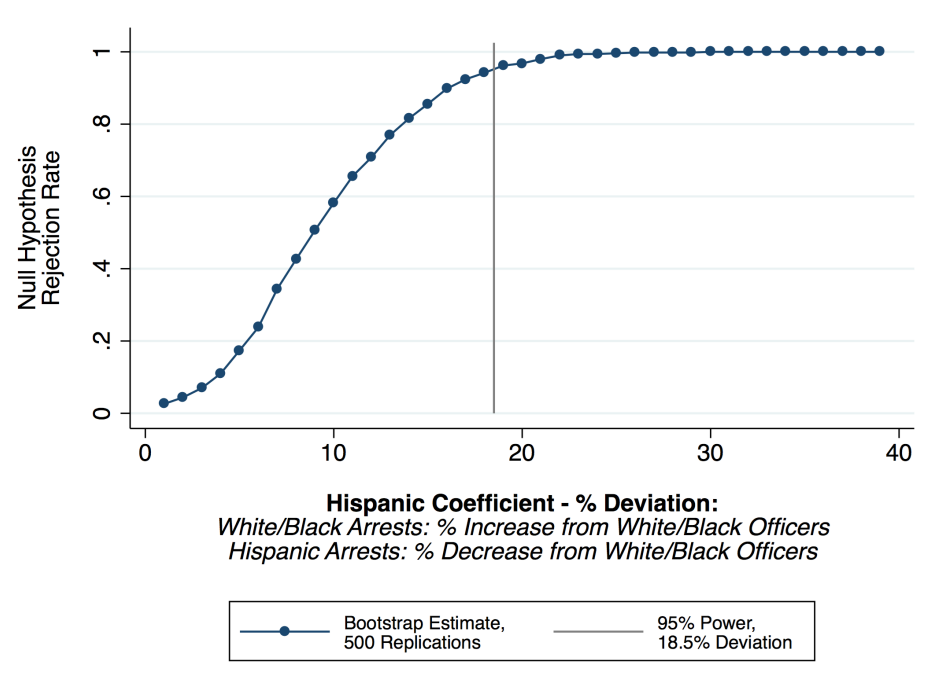
Figure A.3: Power of Racial Bias Test, Bootstrap Simulation

Figure A.3.A: Power of the Test, Deviations in Black Officer Race Coefficient



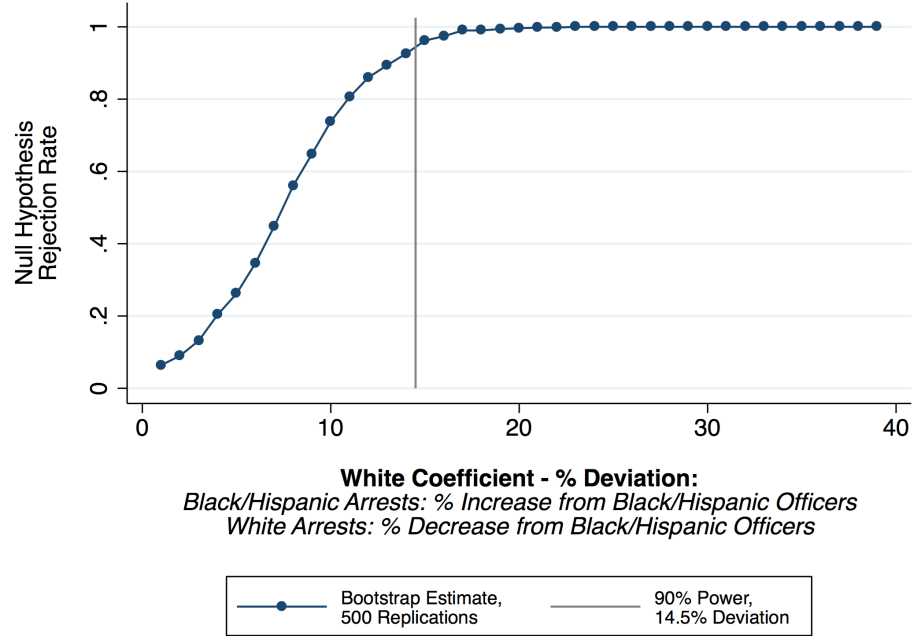
This figure shows results from a bootstrap simulation exercise that estimates the power of the racial bias test used in the paper. This graph plots the rejection rate of the racial bias test when an alternative hypothesis of racial bias is imposed, or the power of the test. The X-axis portrays different alternative hypotheses, expressed as a constant percentage deviation in arrest race outcomes caused by the altered officer race coefficient, relative to the arrest average for officers of other races. The deviation is an increase for arrestee outcomes when the officer and arrestee are different races, and a decrease when the officer and arrestee are the same race.

Figure A.3.B: Power of the Test, Deviations in Hispanic Officer Race Coefficient



This figure shows results from a bootstrap simulation exercise that estimates the power of the racial bias test used in the paper. This graph plots the rejection rate of the racial bias test when an alternative hypothesis of racial bias is imposed, or the power of the test. The X-axis portrays different alternative hypotheses, expressed as a constant percentage deviation in arrest race outcomes caused by the altered officer race coefficient, relative to the arrest average for officers of other races. The deviation is an increase for arrestee outcomes when the officer and arrestee are different races, and a decrease when the officer and arrestee are the same race.

Figure A.3.C: Power of the Test, Deviations in White Officer Race Coefficient



This figure shows results from a bootstrap simulation exercise that estimates the power of the racial bias test used in the paper. This graph plots the rejection rate of the racial bias test when an alternative hypothesis of racial bias is imposed, or the power of the test. The X-axis portrays different alternative hypotheses, expressed as a constant percentage deviation in arrest race outcomes caused by the altered officer race coefficient, relative to the arrest average for officers of other races. The deviation is an increase for arrestee outcomes when the officer and arrestee are different races, and a decrease when the officer and arrestee are the same race.

This simulation is conducted using the following steps (for the White coefficient example). Allow each officer race coefficient in the second stage to be denoted by α_r . I first regress the second stage $\theta_{i,r}$ outcomes on officer demographics excluding the White officer race coefficient, $White_{i,r}$, and recover a predicted $\hat{\theta}'_{i,r}$ estimate and a predicted residual $\hat{r}_{i,r}$. In each bootstrap iteration, I draw a wild bootstrap weight $w_b \in \{-1, 1\}$ with equal probability for each weight. I then impose an alternative hypothesis on the $White_{i,r}$ coefficient to construct a simulated value for each officer effect, $\tilde{\theta}_{i,r}^b = \hat{\theta}'_{i,r} + \Delta White_{i,r} + w_b \hat{r}_{i,r}$. I set the magnitude of Δ to be a constant percent increase for Black and Hispanic arrestee outcomes and an equivalent percent decrease in the White arrestee outcome, so that the alternative hypothesis is a true difference in White officer ranking across the arrestee race outcomes. These percent changes are set relative to Black or Hispanic officer averages for the total arrest outcome, using the higher of the two officer groups (Black or Hispanic) for percent increases and the lower of the two for the percent decrease. In other words, a Δ of a 5 percent deviation will equal the $\Delta = \alpha_{Black} + 0.05 / (\text{mean}(\text{ArrestHispanic}) + \alpha_{Black})$ for the $\theta_{i,Hispanic}$ regression if Black officers were the reference category with larger coefficient, and $\Delta = \alpha_{Hispanic} - 0.05 / (\text{mean}(\text{ArrestWhite}) + \alpha_{Hispanic})$ for the $\theta_{i,White}$ regression if Hispanic officers were the reference category with the smaller coefficient. Using these simulated bootstrap values, $\tilde{\theta}_{i,r}^b$, I regress officer effects on the full set of officer demographic variables and use F-tests to determine whether the ranking of officer race groups changes across arrestee race outcomes. The test is rejected when F-tests are significant at the 10% level and show different officer race rankings across the arrestee race outcomes.

Table A.7: Replication of Racial Bias Tests in Prior Literature

A. Population Benchmark Tests				
	Total	Black	Hispanic	White
Dallas Population	1,260,688	24.1%	41.7%	29.4%
Proportion of Arrests to Total Arrests	9,033	54.7%	24.1%	21.2%
<i>Arrest Rates Fit Population Distribution</i>				
Chi-Squared Statistic	4373			
P-Value	0			
B. Knowles, Persico & Todd (2001)				
	Total	Black	Hispanic	White
Arrests as a Fraction of Suspects	50.2%	48.9%	49.5%	54.5%
<i>Conditional Arrest Rates Equal</i>				
Chi-Squared Statistic	37.13			
P-Value	0			
C. Anwar & Fang (2006)				
	Black Officers	Hispanic Officers	White Officers	
Arrests of Black Suspects	49.0%	50.0%	48.6%	
Arrests of Hispanic Suspects	51.2%	50.2%	48.7%	
Arrests of White Suspects	56.3%	55.4%	53.7%	
<i>Rank of Arrest Rates across Officer Race Independent of Suspect Race</i>				
<i>Officer Arrest Rate Order for each Suspect Race:</i>				
<i>Black ≥ Hispanic ≥ White</i>				
Total Number of Tests	6			
Number of Tests that Reject the Null Hypothesis	0			
D. Antonovics & Knight (2009)				
	Black Officers	Hispanic Officers	White Officers	
Race Match Arrest Rate as a Fraction of Suspects	50.2%	53.3%	54.5%	
Race Mis-Match Arrest Rate as a Fraction of Suspects	50.0%	50.9%	50.4%	
<i>Arrest Rates Equal:</i>				
<i>Officer/Suspect Match and Mismatch (Race or Gender)</i>				
Chi-Squared Statistic	1.229	2.71	16.68	
P-Value	0.268	0.100	0	

This table shows the results of racial bias tests used in the prior literature when applied to the data in this study. Conditional arrest rates are defined as number of incidents with arrests divided by number of incidents with suspects. Each test measures differences in raw aggregated statistics in the data, that are not adjusted for contextual factors related to incidents through regression. Panel (A) is calculated at the incident level, with suspect race categorized as a maximum value if there are multiple suspects. Panel (C) uses a maximum value to categorize officer race at the incident level. Panel (B) and (D) allow for multiple arrestee, suspect, or officer races per incident. In this way, the tests shown here are simplified versions of tests in the prior literature, because in some of the source papers, the authors also conduct tests that adjust for other factors related to incidents in addition to race. The tests only use information prior to 2017, given the availability of the suspect outcome.

A.2 Coefficients in the First Stage of Model

In the body of the paper, I restrict attention to aspects of officer effects because this paper focuses on estimating differences in officer arrest behavior, $\hat{\theta}_i$, and the importance of officers in predicting arrests. This appendix discusses other components of the arrest prediction model.

Table [A.8](#) shows the first stage coefficients for incident characteristics, X_{kt} . First, the probability of an arrest is increasing in call severity or urgency, at a decreasing rate. This is shown by the "Time to Dispatch" variables that measure the number of minutes that lapse between when a call is made by the complainant and when an officer is dispatched to the scene. An increase of 10 minutes in this time gap decreases the likelihood that an arrest is made by 4%. The average call in the data has a time difference of 24 minutes between the call and dispatch time, which corresponds to a decrease the likelihood of arrest by 9% relative to an instantaneously dispatched call.

Next, the model includes direct controls for complainant demographic information, relative to the omitted category of complainants with no demographic data. The likelihood of arrest increases when there are more complainants, and is largest if the complainant is Black. All race categories can be included as controls because some calls have multiple complainants of different races. Arrests are no more likely when the complainant is female. Lastly, arrests are less likely when there is demographic information for the complainant and when there is a victim injury. Complainants without listed

demographics are often businesses, so the negative complainant demographic information coefficient suggests that arrests are more likely when offenses occur in business establishments.

The third set of incident controls in the model are dispatch and location type codes. These variables are generally more positive for crimes that are more serious or are likely to have the evidence necessary to make an arrest. Relative to minor incidents (other minor), arrests are 7 percentage points higher for criminal assaults, shootings, or armed encounters. At the same time, while a robbery is a serious violent offense, these incidents have a lower likelihood of arrest, possibly because suspects are difficult to identify in these incidents. Likewise, burglaries and thefts are less likely to result in arrest than the omitted category. For location, incidents that occur in a business setting appear more likely to result in arrest, while crimes that occur on the street are the least likely to involve an arrest. This may be related to security surveillance systems used in businesses.

Lastly, the model includes indicator variables for the hour within each shift. Interestingly, arrests are less likely in the sixth and seventh shift hour than in the last hour of a shift. This likely relates to officer overtime pay incentives. If officers make arrests in the last hour of their shift, they are more likely to receive overtime pay for activities related to filing the arrest, including booking the individual in the county jail and writing the arrest report for the incident. Officers are 10% more likely to make arrests in the last hour of their

shift than the second to last hour of their shift, relative to the average arrest rate in the sample.

Overall, the incident context controls are important predictors of whether an incident results in an arrest. As noted above, these variables collectively account for $\approx 60 - 70\%$ of the explainable variation in arrest outcomes.

A.3 Empirical Bayes Shrinkage Estimates

As outlined in the text, the estimates of permanent officer arrest propensity are adjusted using Empirical Bayes techniques. Empirical Bayes techniques are useful when a statistician observes a large number of different estimates of parameters that are drawn from the same underlying distribution, and each estimate is measured with error. These techniques are detailed in work by [79], and are commonly employed in the economics of education literature on teacher value added [e.g. 49, 66, 16, 62, 1]. A number of different variants of Empirical Bayes techniques have been used in the prior literature, the estimation in this paper shares features with [49, 16, 1]. In robustness checks in the paper, I show that the results do not change when a number of alternate precision adjustments are used.

In this paper, I observe sample estimates of officer arrest propensity, \bar{r}_i , which are derived from a first stage regression model. Each of these estimates is an approximation of a “true” officer arrest propensity, θ_i , though some officer estimates are derived from more observations and are thus more precise than others. The underlying parameters, θ_i , can also be thought of as random variables which are derived from a separate distribution of potential officer arrest propensities. Empirical Bayes techniques develop a “prior” distribution for the underlying distribution of θ_i that is estimated empirically from the data on all officers. The estimation constructs a weighted mean of the observational estimate and the “prior.” The underlying distribution of θ_i

is known to be centered at 0, given that officer arrest propensity is defined in relative terms and the first stage model includes an intercept.

Each θ_i is assumed to be independent and identically distributed across G total officers. The underlying distribution of each \bar{r}_i and the total distribution of θ_i across i are given by:

$$\begin{aligned}\bar{r}_i|\theta_i &\sim N(\theta_i, \frac{\sigma_{\varepsilon,i}^2}{N_i}) \\ \theta_i|\mu, \sigma_A^2 &\sim N(0, \sigma_A^2)\end{aligned}$$

The mean of the distribution of θ_i is known to be 0 in this setting, given the normalization of the fixed effects in the model. Given a “prior” for the distribution of θ_i , the posterior distribution of $\theta_i|\bar{r}_i$ give the adjusted estimates of $\hat{\theta}_i^{EB}$ used in this paper:

$$\begin{aligned}\theta_i^{EB}|\bar{r}_i, \sigma_{\varepsilon,i}^2, \sigma_A^2 &\sim N(B\bar{r}_i, B\frac{\sigma_{\varepsilon,i}^2}{N_i}) \\ \text{where } B &= \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma_{\varepsilon,i}^2}{N_i}}\end{aligned}$$

I derive estimates of officer arrest propensity, $\hat{\theta}_i^{EB}$, using the following steps:

1. Estimate the first stage of the model and calculate \hat{r}_{ikgt} and \bar{r}_i . I include all officer fixed effects in the first stage regression to allow for arbitrary correlations between responding officers and the other covariates in the model to improve the estimation of the residuals. This procedure is similar to the first stage approach used in [16].

$$\begin{aligned} Arrest_{ikgt} &= \theta_i + \theta_j + \pi X_{kt} + \delta_{gt} + \phi_g + \varepsilon_{ikgt} \\ \hat{r}_{ikgt} &= \hat{\theta}_i + \hat{\varepsilon}_{ikgt} \\ \bar{r}_i &= \frac{1}{N_i} \sum_{N_i} \hat{r}_{ikgt} \end{aligned}$$

2. Calculate individual variance estimates, $\hat{\sigma}_{\varepsilon,i}^2$ and solve for a sample analog of the prior variance of θ_i , $\hat{\sigma}_A^2$.

$$\begin{aligned} \hat{\sigma}_{\varepsilon,i}^2 &= \frac{1}{N_i - 1} \sum_{N_i} (\hat{r}_{ikgt} - \bar{r}_i)^2 \\ \sigma_A^2 &= E[r_{ikgt}^2] - E[\varepsilon_{ikgt}^2] \\ \hat{\sigma}_A^2 &= \frac{1}{N - G - K} \sum_G \sum_{N_i} \hat{r}_{ikgt}^2 - \frac{1}{N - G} \sum_G N_i \hat{\sigma}_{\varepsilon,i}^2 \end{aligned}$$

with $N - G - K$ are the degrees of freedom in the first stage regression, given G officers and K regressors in the first stage model.¹

¹Given that I absorb four sets of fixed effects in the model, θ_i , θ_j , δ_{gt} , and ϕ_g , K contains the number of group categories in the non-focal fixed effects. In practice, the degrees of freedom must also be adjusted for the number of omitted reference categories in the model, or the number of "mobility groups", M . M is estimated to be 9 in this setting. The actual degrees of freedom used is $N - G - K + M$.

3. Calculate the posterior estimates $\hat{\theta}_i^{EB}$ by applying the shrinkage factor \hat{B} . The shrinkage factor is always less than 1 and is increasing in N_i and decreasing in $\sigma_{\varepsilon,i}^2$. This factor gives higher weight to police officer arrest propensity estimates that are more precisely measured and shrinks less precise estimates toward 0, the center of the distribution.

$$\hat{\theta}_i^{EB} = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \frac{\hat{\sigma}_{\varepsilon,i}^2}{N_i}} \cdot \bar{r}_i$$

The following figure displays the relationship between the unadjusted and adjusted estimates of officer effects and the number of observations per officer:

Figure A.4: Adjusted and Unadjusted Officer Effects

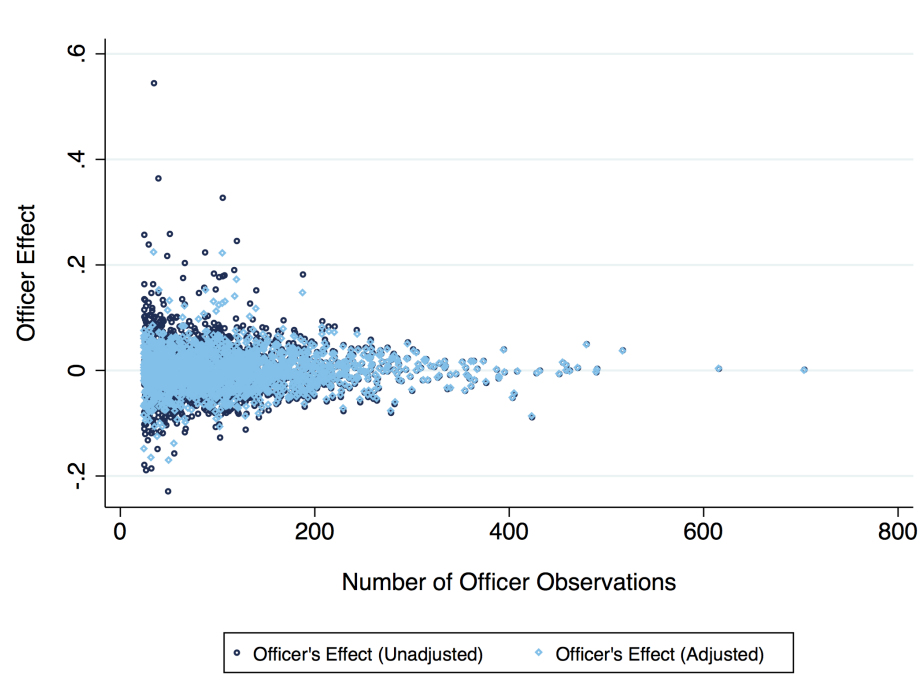


Figure shows unadjusted officer fixed effects overlaid with the adjusted officer effects estimates used in this paper. The correlation between the adjusted and unadjusted estimates is 0.971.

A.4 Economic Model for Racial Bias Test

In this section, I outline a model used to test for the presence of racial bias among officers. The model is adapted from the test of racial bias in [6] to the setting of police calls for service, where the econometrician does not directly observe officer effort choices or suspect race.

Relative Ranking of Arrest Rates by Officer Race

There are two races for officers and suspects in the model, $r \in \{M, W\}$. The model examines officer responses to incidents of a similar type, with the same observable characteristics. When officers arrive to respond to an incident they observe the suspect's race but this information is not observed by the econometrician.² The likelihood that a call has a suspect of a given race, r_s , is ψ^{r_s} . Note that $\psi^M + \psi^W \leq 1$ if some calls for service do not have any relevant suspect, which may occur for accidents or incidents that do not have a clear party that is at fault.

For each suspect race, r_s , π^{r_s} is the likelihood that an arrest is *feasible* if an officer exerts effort to respond. An arrest is feasible if there is a minimum basis for an arrest. Feasibility is a function of the total amount of

²The data does include some information about suspects. The data includes records of suspects identified by officers prior to 2017, as well as characteristics of suspects unknown to officers at the conclusion of a response. However, suspect information is always recorded by officers as part of a response to a call, so it is treated as an outcome of a response rather than a given characteristic of an incident.

evidence that is capable of being recovered for an incident, which may vary according to suspect characteristics that may be correlated with suspect race. Alternatively, the basis for an arrest could be higher if a suspect has a criminal history that may be discovered during a response. Criminal history is a suspect characteristic that may also be correlated with suspect race.

More severe offenses will also have a higher basis for an arrest. The regression application of this test will directly control for observable components of the severity or arrest feasibility of incidents. However, there may be unobservable characteristics of incidents that affect arrest feasibility. If the composition of unobservable characteristics is correlated with suspect race, π^{rs} will also vary across suspect race.

When an officer responds to an incident, he observes information related to the offense that provides a signal of whether an arrest is feasible. This information may include evidence immediately available at the scene, cooperation of the victim, location of the incident, etc. The information at the scene is summarized by an index $s \in [0, 1]$. If an arrest is feasible, s is randomly drawn from the distribution $f_a^{rs}(s)$, while if the arrest is not feasible, s is randomly drawn from the distribution $f_n^{rs}(s)$. These distributions are allowed to differ across suspect race, r_s , reflecting the fact that total information content in responses may differ for different suspect races.

The distributions $f_a^{rs}(s)$ and $f_n^{rs}(s)$ have the following properties:

- Both are defined over the full support of $s \in [0, 1]$
- Monotone Likelihood Ratio Property: $\frac{f_a^{rs}(s)}{f_n^{rs}(s)}$ is strictly increasing in s . This implies that a higher s means an arrest is more likely to be feasible.
- Unbounded Likelihood Ratio: $\frac{f_a^{rs}(s)}{f_n^{rs}(s)} \rightarrow \infty$ as $s \rightarrow 1$. This implies that very high signals θ provide nearly certain information that an arrest is feasible.
- $F_a^{rs}(s)$ first order stochastically dominates $F_n^{rs}(s)$.

Officers make a discrete effort choice $E \in \{0, 1\}$ after viewing the suspect race and the signal, $\{r_s, s\}$. If an officer chooses to exert effort, the posterior likelihood of arrest is increasing in s and is given by Bayes' Rule:

$$P(A|r_s, s) = \frac{\pi^{r_s} f_a^{rs}(s)}{\pi^{r_s} f_a^{rs}(s) + (1 - \pi^{r_s}) f_n^{rs}(s)}$$

Each officer will receive a benefit if an arrest is made that is normalized to 1 and faces a cost of effort that varies by both suspect race and officer race, $t(r_s, r_p) \in [0, 1]$. If the officer chooses not to exert effort, he receives a benefit of zero. Each officer maximizes his utility as a choice between effort and no effort:

$$\max\{P(A|r_s, s) - t(r_s, r_p), 0\}$$

Officers will exert effort in response to an incident when $P(A|r_s, s) \geq t(r_s, r_p)$. As a result, it can be shown that officers will exert effort on a suspect of race, r_s , if the value of $s \geq s^*(r_s, r_p)$, where the threshold $s^*(r_s, r_p)$ satisfies $P(A|r_s, s^*(r_s, r_p)) = t(r_s, r_p)$. This effort threshold is monotonically increasing in $t(r_s, r_p)$.

In the setting of calls for service, both officer effort choices and suspect race are not observed in the data. Instead, the econometrician can observe the number of arrestees of a given race adjusted by the total number of incidents in the data, or “unconditional” arrestee race outcomes. Allow these unconditional arrestee race outcomes to be denoted as the “arrest rate” for an arrestee race group and officer race combination.

In terms of the model, the arrestee race outcome for arrestees of race, r_s , and officers of race, r_p , is:

$$A(r_s, r_p) = \psi^{r_s} \pi^{r_s} [1 - F_a^{r_s}(s^*(r_s, r_p))]$$

This arrest rate is decreasing in the signal threshold $s^*(r_s, r_p)$ and is also decreasing in the cost of effort $t(r_s, r_p)$.

The following definitions characterize officer race specific costs:

1. Racial Bias: Officers are racially biased with respect to suspects if for some officer race, r_p , $t(M, r_p) \neq t(W, r_p)$.

2. Monolithic Behavior: Officers are not monolithic in their behavior if officer costs differ across officer race for a given suspect race, or $t(r_s, M) \neq t(r_s, W)$.
3. Statistical Discrimination: Assume $t(M, r_p) = t(W, r_p)$, or officers are not racially biased. Then race r_p officers will exhibit statistical discrimination if $s^*(M, r_p) \neq s^*(W, r_p)$.

If officers are not racially biased and exhibit monolithic behavior across officer race, then $t(M, M) = t(M, W) = t(W, W) = t(W, M)$. It follows that arrest rates within suspect race will be constant across officer race, but that total arrest rates for different suspect races may differ if $s^*(M, r_p) \neq s^*(W, r_p)$, or there is statistical discrimination.

If officers do not exhibit monolithic behavior but are also not prejudiced, then the ranking of arrest rates across officer race within suspect race will be independent of suspect race, or constant across suspect race.

For example, allow minority officers to have a higher cost of effort than White officers for any race of suspect. Then:

$$\begin{aligned}
t(M, M) &> t(M, W) \quad \& \quad t(W, M) > t(W, W) \\
t(M, M) &= t(W, M) \quad \& \quad t(M, W) = t(W, W) \\
s^*(M, M) &> s^*(M, W) \\
\& \quad s^*(W, M) &> s^*(W, W) \\
A(M, M) &< A(M, W) \\
\& \quad A(W, M) &< A(W, W)
\end{aligned}$$

Or in this case, minority officers will be less likely to make arrests than White officers for both suspect races. In other words, the relative ranking of minority and White officers is the same for both suspect groups. We can conclude that if both races of officers are not biased, the relative ranking of arrest rates across officer race should be the same for incidents within each suspect race.

Generally, the test proposed in this paper allows total arrest rates to differ across arrestee race by focusing attention on relative rankings of officer arrest rates rather than total levels of officer arrest rates. This feature allows officers to behave in a manner that is consistent with statistical discrimination and isolates officer behavioral patterns associated with taste-based racial bias. Statistical discrimination will occur in this model if total arrest rates for one suspect group is higher than the other suspect group but the relative ranking of officer arrest rates is the same for both suspect groups. For example, it may be the case that minority suspects are more likely to have a criminal

history and this causes the total signal threshold to be lower for incidents with minority suspects, $s^*(M, r_p) < s^*(W, r_p)$. In the example above, this will occur if both White and minority officers make more arrests of minority suspects than White suspects, $A(M, M) > A(W, M)$ and $A(M, W) > A(W, W)$, but White officers always arrest suspects at higher rates than minority officers, $A(M, W) > A(M, M)$ and $A(W, W) > A(W, M)$.

If arrests are higher for White officers relative to minority officers when responding to incidents with minority suspects, $A(W, W) < A(W, M)$, and arrests are higher for minority officers relative to White officers when responding to White suspects, $A(M, M) < A(M, W)$, we can conclude that one or both officer race groups is biased. This opposing rank order violates the null hypothesis of no racial bias among officers.

This is illustrated by the following stylized example:

$$\begin{aligned}
& t(M, M) > t(W, M) \quad \& \quad t(W, W) > t(M, W) \\
& \& \quad t(W, W) = t(M, M) \quad \& \quad t(W, M) = t(M, W) \\
& \quad s^*(M, M) > s^*(M, W) \\
& \quad \& \quad s^*(W, M) < s^*(W, W) \\
& \quad A(M, M) < A(M, W) \\
& \quad \& \quad A(W, M) > A(W, W)
\end{aligned}$$

[6] show that the test will effectively identify racial bias when $[t(W, W) - t(W, M)][t(M, W) - t(M, M)] < 0$, or officers of different races have opposing

cost differences for different suspect races. The test will fail to identify racial bias when $[t(W, W) - t(W, M)][t(M, W) - t(M, M)] > 0$, or one officer race group has higher costs for all suspect groups than the other officer race group.

The test also allows officers to have differing total costs of effort that vary by officer race, or behave in a manner that is not monolithic. The first half of this paper assesses whether individual officers behave differently from one another in their responses to similar incidents, which can be interpreted as evidence that individual officers are not monolithic in their behavior.

A.5 Data Appendix

Several different data files were used for this project. This Appendix summarizes the decisions made in cleaning and constructing the data set used for this project.

Incident Data The base file used in this project is the DPD “Police Incidents” file accessed through the Dallas Open Data portal. Data sets compiled and released through this portal are updated daily, with each new data set consisting of a moving time window of records. Because old records are replaced with new records in this interface, I have periodically downloaded new versions of the data (on an approximate monthly basis), updating the existing records with new downloads. Throughout this project, I use the most recently updated record for each incident in the data when there are duplicate records for the same incident across downloaded versions. I allow a grace period of one month to pass before using an incident record, as records may be corrected retrospectively. This procedure allows me to use the most complete set of information for each observation and increase fidelity in comparisons of incidents over time, as some records may be collected, modified, or updated retrospectively. This data updating process is important because DPD releases the records as they occur and some records may be incomplete or inaccurate. The data in this project covers the time window of December 2014 through mid-October 2017.

The data includes all incidents that are reported to DPD by a reporting party or complainant, with the exception of sexually oriented offenses, offenses involving juveniles, and social service referral offenses. While the data encompasses calls for service, it also includes other incidents reported by complainants through other means, as well as some officer-initiated interactions. I use information included in several of the data fields to narrow the set of observations to include only records that are both “highly likely” to correspond to calls for service and are relatively complete.

First, I exclude observations that do not have information on the time that a call was received or dispatched. Second, I exclude calls that do not have a listed police division where the call occurred or have a missing address.

Third, I exclude incidents that were dispatched more than 2 hours after a call was received. I am unable to distinguish between 9-1-1 calls and 3-1-1 calls in the data. While 3-1-1 calls are placed for less serious incidents than 9-1-1 calls, they are connected to the same call-taker lines as emergency 9-1-1 calls and follow the same protocols for response if a response is warranted. Because I only consider calls that received a response and were dispatched within 2 hours that the call was received, the observations are weighted toward 9-1-1 calls. Additionally, by including a control for the urgency of a call in the estimation model I am able to generally distinguish between these two types of calls in the model.

Fourth, I exclude any incident that does not include a complainant

record, or that has a complainant listed as the City of Dallas, the Dallas Police Department, or another police department or sheriff, as these calls are unlikely to involve a civilian complainant. Next, through conversations with officers and dispatchers at DPD, I identify a set of dispatch codes that were unlikely to originate with a call from a complainant. These include calls where an officer is not the first responder but is called to join another officer on an existing call, such as assisting another officer that needs help or is injured, assisting in a chase or foot pursuit, assisting an off-duty officer, providing warrant service for an individual interacting with police, and responding to a fire or aiding a fire department response at their request. Further, I do not include calls to respond to a suspect operating a car that was planted or is being monitored by DPD (“bait” car or ETS activation responses). I also exclude calls that originate with complainants walking into a police station to alert police about an incident as well as 9-1-1 hang-up calls. Lastly, I exclude calls where an officer may have discovered an incident or complainant during patrol that was unlikely to be called in and dispatched to a larger set of officers. These dispatched call types are routine investigations, traffic stops, and public park checks.

To clean the data, I conduct the following steps. I calculate the difference between the time a call is made and the time that the call is dispatched in minutes to form a call severity or call urgency variable, using time stamps in the data. In a very small number cases where there are negative time differences, I assume that these time stamps were mistakenly swapped

and take the absolute value of the difference. I follow a similar procedure to calculate the time of an officer response after dispatch, but replace negative values as zeros in this case. While there are codes for shifts or “watches” in the data, these often do not align with the general time slots for shifts. I construct a more strict and usable definition of shifts to eliminate shift overlap; these shifts are 12am-8am, 8am-4pm, and 4pm-12am. Next, I combine the remaining 116 dispatch codes into 15 groupings, to increase power and remove very small categories. Similarly, I combine the 34 location type codes in the data into 11 groupings.

Throughout the analysis, I use dispatch codes as incident controls rather than offense types, because dispatch codes are available to officers before they respond to incidents while incident offense types are designated by patrol officers after they arrive at the scene of an incident, and therefore are a choice variable.

Persons Involved and Arrest Data I use three additional DPD open data files to supplement information on arrests and complainants in the main “Police Incidents” file: “Police Arrests”, “Police Arrest Charges”, and “Police Person”. These files offer different coverage of arrests than the “Police Incidents.”, and I use the differences across the files to create a comprehensive measure of arrests. These files contain more detailed records of arrestees, including the time and place of arrest, arrestee name and demographics, and arrest charge. The “Police Person” data includes names and demographic information for

arrestees as well as suspects and complainants associated with incidents (but is only available through 2017). The “Police Incident” data includes demographic information for complainants and badge identifiers and names for officers. The data contains identifiers that allow records to be merged across files.

The main outcome in this paper is whether any arrest occurred in association with a particular incident. I consider an arrest to have occurred if there is a record of an arrest in the “Police Incidents” file or any of the three supplementary arrests files. Because the supplementary arrest files are typically compiled after the incident file and include more detailed information about arrestees, I consider an arrest as being verified if there is a record of the arrest in any of the supplementary files.

I use complainant records in the “Police Person” to supplement the complainant records in the main incident file, and use these combined records to exclude records that do not have a civilian complainant. When there are multiple complainants, arrestees or suspects, associated with a call, demographic characteristics are measured as the maximum of these variables across the relevant group. For example, the indicator for a Black complainant is set to 1 if any of the complainants associated with the call are Black. I calculate whether there was a victim injury in a similar way, using information on complainants from both the “Police Person” and incident files.

I determine whether an arrest was for a felony or misdemeanor offense using arrest charge and offense severity codes in both the “Incident”

and “Police Arrest Charges” file. Incidents may have both a felony and misdemeanor arrest if a suspect was arrested for multiple offenses or there is more than one arrestee associated with an incident.

Officer Demographics Data Lastly, I complement the open data files available from DPD with officer information obtained through a Freedom of Information Act (FOIA) request to the city of Dallas. Through this FOIA request, I acquired records of all police department employees from 2014 to the present that include officer names, badge number, job title, hire date, leave date (if applicable), ethnicity or race, gender, age, and salary. I match the FOIA request records to the incident file using officer badge numbers, and match officers by name if badge numbers are not available.

Because the FOIA request file includes employee title and badge number, I also use this information to exclude observations with responding officers that are civilian police employees, as these incidents likely involved only a phone response and did not entail a physical patrol officer response to the scene. I also exclude observations for individuals that are not employed by DPD (e.g. county police, firefighters, volunteer police officers).

Appendix B

Appendix:

Safety in Police Numbers:

Evidence of Police Effectiveness from COPS Grant Applications

B.1 Appendix Tables and Figures

Table B.1: Impact of Police on Crime, Influence of Model Type and Covariates

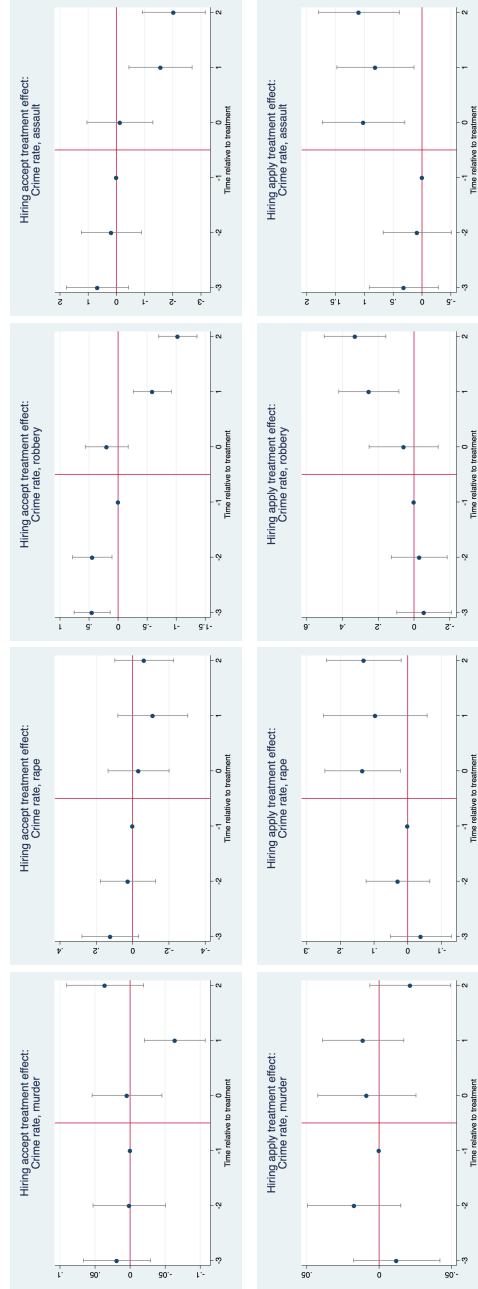
	Violent Crime				Property Crime			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Police	0.752*** (0.065)	0.542*** (0.060)	0.265*** (0.048)	0.263*** (0.048)	6.405*** (0.446)	5.933*** (0.467)	2.205*** (0.403)	2.173*** (0.405)
Elasticity	0.508	0.366	0.179	0.178	0.474	0.439	0.163	0.161
First Stage: Police Rate								
Accept Hiring	1.586*** (0.269)	1.199*** (0.247)	0.636*** (0.072)	0.648*** (0.073)	1.56*** (0.27)	1.19*** (0.248)	0.646*** (0.072)	0.658*** (0.073)
F Test: Accept Hiring	34.88	23.64	77.84	79.63	33.48	23.07	80.32	82.23
% Effect: Accept Hiring	6.74%	5.10%	2.70%	2.75%	6.63%	5.06%	2.74%	2.80%
Reduced Form								
Accept Hiring	17.88*** (0.946)	14.92*** (0.951)	-1.428** (0.456)	-1.227** (0.461)	78.68*** (5.01)	62.91*** (4.997)	-7.433** (2.308)	-6.52** (2.331)
% Effect: Accept Hiring	51.4%	40.8%	-4.10%	-3.53%	24.7%	19.8%	-2.34%	-2.05%
IV								
Police	11.27*** (1.801)	12.438*** (2.572)	-2.244** (0.768)	-1.892* (0.75)	50.45*** (8.175)	52.86*** (10.8)	-11.5** (3.92)	-9.904** (3.808)
Elasticity	7.62	8.41	-1.517	-1.280	3.732	3.911	-0.851	-0.733
Year FE	X	X	X		X	X	X	
Year X City Size FE				X				X
Covariates		X	X	X		X	X	X
Police Department FE			X	X			X	X
Police Rate Mean	23.53	23.53	23.53	23.53	23.54	23.54	23.54	23.54
Crime Rate Mean	34.8	34.8	34.8	34.8	318.1	318.1	318.1	318.1
Observations	93,081	93,081	93,081	93,081	93,296	93,296	93,296	93,296
Number of Departments	6,966	6,966	6,966	6,966	6,964	6,964	6,964	6,964

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level in models with police department fixed effects. Crime outcomes and police variables are per 10,000 residents in a municipality. Each specification successively controls for more variables. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants. The preferred specification throughout the paper corresponds to (4) and (8).

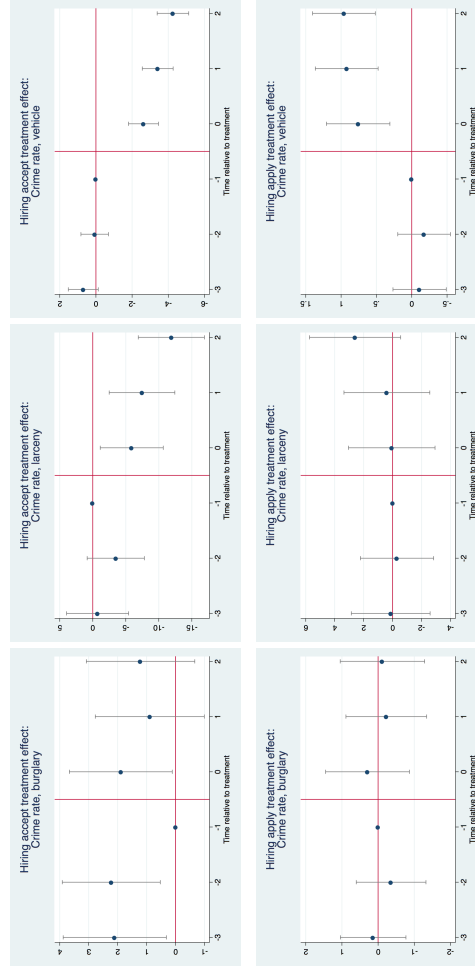
Figure B.1: Timing of Treatment, Specific Crimes, Application and Acceptances

Figure B.1.A: Violent Crime Categories



The range around each estimate represents a 95% confidence interval using robust standard errors clustered at the police department level from the preferred specification (as in Figure 2.2). Coefficients in the acceptance and application graphs for each outcome are derived from the same regression. Graphs are created by duplicating balanced panels of 6 years corresponding to a centered treatment year for each year between 2000-2012. Treatments are indexed to the first year of a new hiring grant acceptance or application (year 0 above).

Figure B.1.B: Property Crime Categories



The range around each estimate represents a 95% confidence interval using robust standard errors clustered at the police department level from the preferred specification (as in Figure 2.2). Coefficients in the acceptance and application graphs for each outcome are derived from the same regression. Graphs are created by duplicating balanced panels of 6 years corresponding to a centered treatment year for each year between 2000-2012. Treatments are indexed to the first year of a new hiring grant acceptance or application (year 0 above).

Table B.2: Impact of Police on Index I Arrests, By Demographic Group

	(1) Total	(2) White	(3) Black	(4) Male	(5) Age 0-15	(6) Age 15-24	(7) Age 25-39
Murder	-0.024 (0.026)	-0.016 (0.017)	-0.007 (0.017)	-0.021 (0.024)	0.000 (0.003)	0.006 (0.016)	-0.025+ (0.014)
Elasticity	-1.973	-2.664	-1.359	-2.048	0.0724	1.130	-6.160
Y Mean	0.280	0.143	0.128	0.243	0.00358	0.126	0.0964
Rape	0.033 (0.044)	0.019 (0.032)	-0.004 (0.023)	0.027 (0.043)	0.012 (0.011)	0.030 (0.025)	-0.024 (0.020)
Elasticity	0.955	0.753	-0.405	0.782	5.577	1.965	-2.138
Y Mean	0.813	0.575	0.209	0.800	0.0493	0.351	0.259
Robbery	0.026 (0.085)	-0.004 (0.048)	0.025 (0.063)	0.016 (0.077)	0.002 (0.014)	0.076 (0.064)	-0.032 (0.036)
Elasticity	0.240	-0.0780	0.475	0.174	0.576	1.253	-1.064
Y Mean	2.477	1.179	1.242	2.191	0.100	1.407	0.706
Assault	-0.161 (0.346)	-0.060 (0.242)	-0.119 (0.183)	-0.188 (0.274)	-0.072+ (0.043)	-0.128 (0.148)	-0.009 (0.153)
Elasticity	-0.265	-0.143	-0.697	-0.388	-2.720	-0.607	-0.0422
Y Mean	14.18	9.702	3.990	11.28	0.621	4.903	5.118
Observations	81,605	81,605	81,605	81,605	81,605	81,605	81,605
Burglary	0.197 (0.237)	0.201 (0.194)	0.041 (0.110)	0.058 (0.208)	-0.055 (0.082)	0.076 (0.153)	0.056 (0.088)
Elasticity	0.447	0.633	0.376	0.152	-1.209	0.330	0.492
Y Mean	10.27	7.425	2.538	8.916	1.058	5.377	2.640
Larceny	3.215 (3.018)	0.761 (0.810)	-0.585 (0.366)	0.725 (1.534)	-0.519** (0.178)	3.777 (2.894)	-0.001 (0.389)
Elasticity	1.501	0.490	-1.221	0.561	-2.894	3.900	-0.00144
Y Mean	49.96	36.22	11.18	30.12	4.183	22.58	14.17
Vehicle	2.631 (2.841)	-0.075 (0.083)	-0.106* (0.052)	1.190 (1.421)	2.803 (2.841)	-0.136* (0.068)	-0.046 (0.044)
Elasticity	15.98	-0.797	-3.613	9.736	61.30	-2.047	-1.233
Y Mean	3.838	2.197	0.681	2.851	1.066	1.551	0.873
Observations	81,786	81,786	81,786	81,786	81,786	81,786	81,786

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes, arrest outcomes and police variables are per 10,000 residents in a municipality. Arrest rates in each column are total arrests for a given demographic group per 10,000 total residents in a municipality. All models shown have the preferred set of covariates, including demographic covariates, application and acceptance controls for other grants, year by city size fixed effects and police department fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants. Arrest outcomes are not expressed relative to reported crime rates but as total arrests per 100,000 residents. The sample size is different for each arrest type, given the sample cleaning procedure (see Data Appendix C.2).

Table B.3: Impact of Police on Other Arrests, By Demographic Group

	(1) Total	(2) White	(3) Black	(4) Male	(5) Age 0-15	(6) Age 15-24	(7) Age 25-39
<i>Marijuana Sale</i>	0.623** (0.198)	0.374* (0.159)	0.234** (0.076)	0.546** (0.173)	-0.004 (0.017)	0.387** (0.138)	0.216*** (0.062)
Elasticity	3.815	2.845	0.880	3.286	0.0942	2.460	0.917
Y Mean	3.806	3.066	6.199	3.878	-1.058	3.671	5.497
<i>Marijuana Possess</i>	2.249* (0.943)	1.948* (0.824)	0.471 (0.324)	2.014** (0.771)	0.019 (0.050)	1.141* (0.533)	0.682* (0.336)
Elasticity	1.665	1.899	1.822	1.766	0.591	1.403	1.891
Y Mean	31.50	23.92	6.026	26.60	0.757	18.98	8.417
<i>Narcotics Sale</i>	-0.431* (0.192)	-0.107 (0.099)	-0.358** (0.130)	-0.334* (0.153)	-0.009 (0.007)	-0.238** (0.087)	-0.113 (0.082)
Elasticity	-3.984	-2.231	-6.077	-3.765	-10.68	-5.563	-2.584
Y Mean	2.526	1.114	1.373	2.070	0.0194	0.999	1.022
<i>Narcotics Possess</i>	-1.302*** (0.384)	-0.447+ (0.256)	-0.858*** (0.206)	-1.050*** (0.300)	-0.020* (0.009)	-0.668*** (0.158)	-0.461** (0.171)
Elasticity	-4.240	-2.205	-8.600	-4.451	-13.56	-6.695	-3.575
Y Mean	7.163	4.733	2.326	5.500	0.0340	2.327	3.008
Observations	81,423	81,423	81,423	81,423	81,423	81,423	81,423
<i>Simple Assault</i>	-0.151 (0.864)	-0.315 (0.619)	0.012 (0.443)	-0.364 (0.652)	-0.484** (0.170)	-0.102 (0.355)	0.116 (0.345)
Elasticity	-0.0700	-0.202	0.0223	-0.228	-3.504	-0.134	0.147
Y Mean	50.37	36.24	12.38	37.18	3.221	17.76	18.43
<i>DUI</i>	3.371* (1.311)	3.055* (1.187)	0.600+ (0.320)	2.838** (1.058)	-0.000 (0.007)	0.766+ (0.398)	1.291* (0.597)
Elasticity	1.217	1.286	2.600	1.275	-0.369	1.067	1.147
Y Mean	64.55	55.36	5.378	51.86	0.0232	16.73	26.25
Observations	82,438	82,438	82,438	82,438	82,438	82,438	82,438

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Standard errors in each model are robust and clustered at the police department level. Crime outcomes, arrest outcomes and police variables are per 10,000 residents in a municipality. Arrest rates in each column are total arrests for a given demographic group per 10,000 total residents in a municipality. All models shown have the preferred set of covariates, including demographic covariates, application and acceptance controls for other grants, year by city size fixed effects and police department fixed effects. Demographic variables include population, age distribution, racial distribution, proportion male, unemployment rate and average pay. All models control for hiring application and applications and acceptances of other COPS grants. Arrest outcomes are not expressed relative to reported crime rates but as total arrests per 100,000 residents. The sample size is different for each arrest type, given the sample cleaning procedure (see Data Appendix C.2).

B.2 Data Appendix

Grant Data I obtained data on all accepted and rejected grant applications submitted to the COPS Office between 1994-2014 through a Freedom of Information Act (FOIA) request. Both accepted and rejected grant files included information on the agency address, grant program, proposed project start date, and project end date. The accepted grant file additionally included the amount of money awarded, and the number of officers funded by the grant, although this information was not always complete. The 54 grant programs administered in this period were consolidated into groups based on grant purpose: police hiring grants and other grants. Multiple grant acceptances and applications were taken consolidated within police department, year, and grant category cells. If a department had any acceptance within a department-year-category, the observation was coded as an acceptance within that category.

Because police employment data is recorded as a snapshot statistic on October 31st of each year, the start dates of each grant were adjusted such that grants beginning in November and December were indexed to begin in the following year. Within each grant program, the project length of grants was determined by taking the modal project time across accepted grants. Next, to improve the merge between COPS data and the crime data in this project, I manually inspected over 500 grant observations that did not merge to the administrative crime data and corrected ORI police department ids whenever possible.

The COPS grant database includes 3 categories of application status for the municipal police departments in the analysis sample: approved, rejected, and withdrawn. The withdrawn category represents grants that were stopped after being approved and were included within the accepted grant data file compiled in the FOIA request. Approximately 15% of all accepted grants were withdrawn between 1994-2014. I treat withdrawn grants as accepted grants in the analysis, though the results are not sensitive to this designation (see Table 2.6).

Crime, Police and Arrest Data The UCR micro data includes information on a number of different police department types, including state police, county police, campus police and tribal police. In this project, analysis is restricted to municipal police agencies in order to best link the crime data to geographically based demographic data, using information on police department type from the 2005 UCR Crosswalk file. Agencies were included if their government type was classified as "township" or "municipal," or if their agency type was classified as

"municipal police." In geographic units (Census Places) with no municipal police agencies, "sheriffs" were included as municipal police districts. A number of districts were recoded where errors were found, either because of mismatches in codes to the UCR micro data or because of multiple municipal agencies attributed to the same Census Place.

As discussed in the Data Section above, the UCR Offenses Known and Clearances by Arrest contains crime and clearance data is voluntarily submitted by police agencies and is not formally audited. This data has several gaps and likely contains a large number of reporting errors. As a first step in cleaning the data, I identified values that may have been entered to signify missing data even though these values were not specified in the codebooks (e.g. high frequency occurrences of 999, 9999 etc.) and replaced negative reported crime and crime clearance values as missing. Next, I replaced agency year observations as missing if all reported crimes in that year were zero and the agency had high average levels of crime over time, or the average non-zero crime level for the agency was above the median (over all agencies) of agency average crime levels (averaged across years) in any of the 7 violent and property crime categories. Lastly, I marked observations as missing that had a greater number of cleared crimes than prior unsolved crimes for a given agency or that had cleared crimes that were more than ten times the number of reported crimes reported in a category in a given year. After this data was merged to the other data sets used in this paper, an algorithm was used to identify additional outliers and mark these outliers as missing as well (see below for a description). All observations with missing values were omitted from the final analysis sample used in the paper.

The policing data was drawn from the UCR Law Enforcement Officers Killed and Assaulted (LEOKA) database, which includes records of the number of sworn police and civilians employed by police agencies in each year. Before merging this data with the other data sets, police agencies that recorded having zero police officers were coded as missing, as this case is not possible for an agency that is reporting data.

Arrest data was drawn from the UCR Arrest Files. The arrest data is difficult to clean because agencies do not commonly report arrest categories for which they did not arrest any individuals. This makes it tricky to discern whether a missing value is a true zero or is actually missing. To deal with this, I organized commonly reported crimes into groupings, and made the assumption that if an agency reported one type of a crime within a crime group and not others within the same group, the missing categories were likely zero. The crime groupings used were the 7 Index violent and property crimes (group 1), common drug crimes (sale and possession of

marijuana, narcotics and other drugs) (group 2), as well as simple assaults and driving under the influence of alcohol or DUIs (group 3). Group 1 includes zeros rather than missing values when any of the categories within group 1 or other groups are positive and non-missing. Groups 2 and 3 include zeros when one of their subcategories is positive and non-missing or there is a positive and non-missing category in group 1. Lastly, the arrest data includes counts of arrests by gender, race and age group. Zero values for these group counts were imputed to match the total arrest categories for each recorded crime.

Demographic Data Demographic data used to create covariate controls in this paper was derived from the Census Bureau and the Bureau of Labor Statistics (BLS). The geographic unit of analysis in this paper is the Census FIPS Place level, which corresponds to the geographic borders of towns, municipalities, and cities. Population counts are available at the Place level in each year from the Census, but all other demographic information is only available at the county level. Counties may contain Places or overlap with them, so a Census data set that includes population counts of individuals in each county within a Place was used to create weighted averages of demographic variables at the county level. The demographic variables included in the Census county files are race, age and gender. Similarly, BLS data on unemployment rates and income levels at the county level were used to construct county weighted measures of these variables at the Census Place level. In some cases, population and other demographic data at the county level was missing in some county-years. This data was interpolated as the average of adjacent years when an observation was missing but data was available for the prior and following years.

Omitting Outliers Authors of academic papers that use the UCR crime and police data have used a number of different procedures to identify outliers and clean the data to exclude outliers. The algorithm used in this paper is based off of the procedure used in [34]. Though only the years 2000-2014 are used for the main analyses in this paper, I cleaned and merged data from 1990 to 2014 for analysis (the longer sample is used as a robustness check). For each police department in the sample with at least 10 years of observations, I ran a separate regression of violent crime, property crime, police officers, violent arrests, property arrests, drug arrests and other arrests (where each outcome is a sum of crimes within a category) on a quartic time trend. The fitted values were then used to identify outliers. In [34], the authors visually inspect each observation where the difference in the observed outcome and the predicted outcome is more than 50% and then correct the data where they can. However, in [34], the authors also restrict their sample to cities with

over 10,000 residents using the rationale that they may have fewer reporting errors. This paper includes all towns with populations over 1,000 residents, and because of this larger sample, identified outliers are not visually inspected or replaced with imputed values. Instead, each identified outlier is replaced as missing and is omitted from the final analysis sample.

Because reported crime values are noisier in smaller towns, the 50% threshold for differences between predicted and actual values is too strict for the smaller towns in the sample. To address this feature of the data, the threshold of exclusion is set to be the maximum of the 99th percentile of percent difference between observed and predicted values of all observations in a population group, or the 50 percent difference. Population groups are defined as 1,000 to 1,999 residents, 2,000 to 4,999 residents, 5,000 to 9,999 residents, 10,000 to 24,999 residents, 25,000 to 49,999 residents, 50,000 to 99,999 residents and 100,000 residents or more. Cities are grouped by using the modal population category of a city across years in the sample. Once the outlier values have been identified and replaced as missing for the violent and property sum crime outcomes, subcategories within each group are marked as missing if the total is missing. This procedure creates a different sample size for violent and property crime outcomes and groupings of arrest outcomes. Over 10% of the analysis sample was identified as outliers for districts with 1,000 to 1,999 residents, while this proportion was approximately 1% for districts with greater than 100,000 residents. This methodology reduces the analysis sample size by approximately 7%.

Appendix C

Appendix:

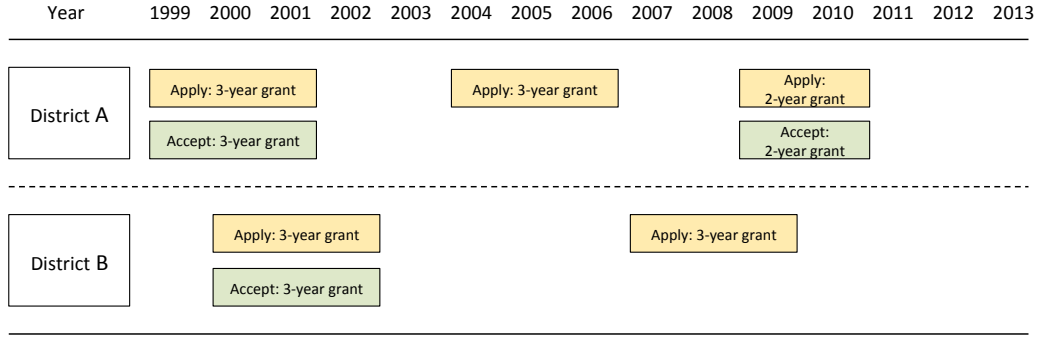
Patrolling Public Schools:

The Impact of Funding for School Police on Student Discipline and Long-term Education

Outcomes

C.1 Appendix Tables and Figures

Figure C.1: Depiction of Model Identification, Comparison of Two Hypothetical Districts

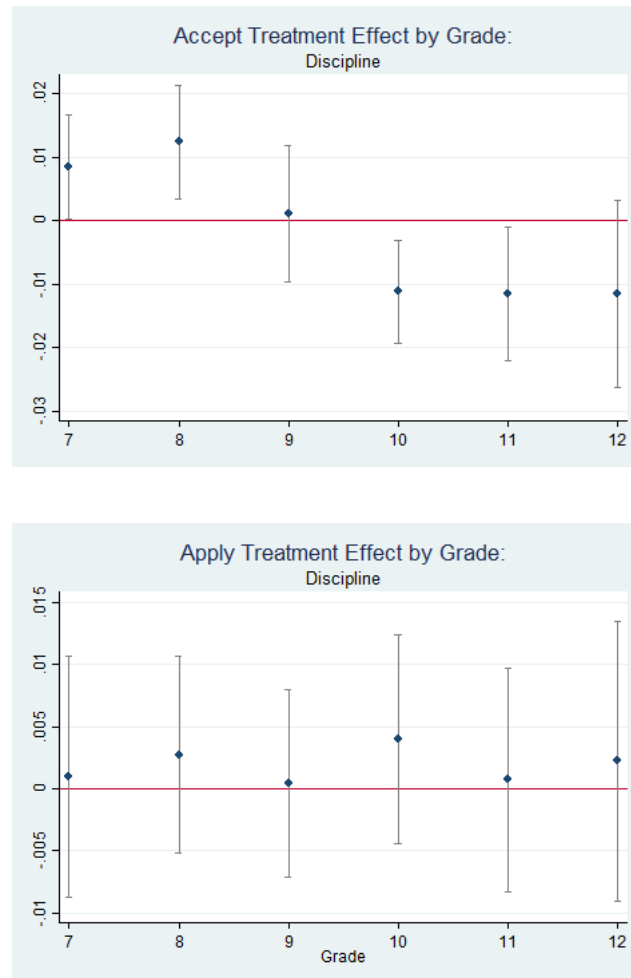


$$\begin{aligned}
 ShortTermOutcome_{igdt} = & \beta_{1m} Accept_{dt} * MiddleSchool_{gt} + \beta_{2m} Apply_{dt} * MiddleSchool_{gt} \\
 & + \beta_{1h} Accept_{dt} * HighSchool_{gt} + \beta_{2h} Apply_{dt} * HighSchool_{gt} \\
 & + \pi X_{igdt} + \delta_t + \gamma_g + \phi_d + \varepsilon_{igdt}
 \end{aligned}$$

$B_{1m,1h}$: This is the conditional acceptance coefficient of interest. It compares “Accept” years to “No Apply” years. For example, this coefficient measures the difference in outcomes between 2001 and 2002 within District A or 1999 and 2000 within District B.

$B_{2m,2h}$: This is the application control. It nets out the effect of district application choices on outcomes. Compares “Apply” years to “No Apply” years. This coefficient also descriptively shows how outcomes change when districts are rejected for grants. For example, difference in outcomes between 2003 and 2004 within District A or 2009 and 2010 within District B.

Figure C.2: Discipline Grant Treatment Effects, by Grade



This figure shows the impact of grant variables split by student grade level. Both figures show coefficients from the same regression, which corresponds to the fully specified model with student covariates, student grade enrollment, and fixed effects for grade, year, and school district. Bars around estimates represent the 95% confidence interval, corresponding to standard errors that are clustered at the school district level.

Table C.1: Grant Effects on School District Budget Data

	(1) Security Expense	(2) Security Expense per Student	(3) Security Expense Ratio	(4) Have a Security Fund
Accept	-182,478 (362,092)	-0.71 (2.29)	0.0001 (0.00008)	0.013** (0.004)
Apply	378,159+ (199,546)	2.50+ (1.35)	-0.00001 (0.00006)	-0.007 (0.004)
Observations	13,498,605	13,497,676	13,498,367	13,498,605
Y Mean	2,679,000	53.0	0.002	0.953
X Mean	0.288	0.288	0.288	0.288
% Effect of Conditional Grant Receipt	-6.8%	-1.3%	5.3%	1.4%

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Robust standard errors clustered at the school district level are shown in parentheses. The suggestive "first stage" regressions above measure budget outcomes at the district-year level for the student-level data sample used in the primary analysis. To make the models comparable to the primary results, the covariates and specification is the same as the preferred specification in the short-term model, with controls for student-district-grade enrollment, student race, gender, economic disadvantage status, special education status, LEP status, and gifted and talented status. Models also include grade and year fixed effects.

Table C.2: Effects by COPS Grant Type

Table C.2.A: Short-term Student Discipline Outcomes

	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
Accept Effects, Middle School				
COPS in Schools (CIS)	0.014** (0.005)	0.007 (0.005)	0.009* (0.004)	-0.0004 (0.0003)
Secure Our Schools (SOS)	-0.017 (0.020)	-0.009 (0.019)	-0.016 (0.019)	-0.0005 (0.0023)
Other Grants	-0.010 (0.010)	-0.010 (0.009)	-0.005 (0.009)	-0.0004 (0.0011)
Accept Effects, High School				
COPS in Schools (CIS)	-0.004 (0.004)	-0.009* (0.005)	-0.002 (0.003)	-0.0010*** (0.0003)
Secure Our Schools (SOS)	0.001 (0.010)	0.004 (0.013)	0.003 (0.005)	-0.0017+ (0.0009)
Other Grants	-0.032* (0.015)	-0.028* (0.013)	-0.003 (0.005)	-0.0007 (0.0010)
Observations	13,596,577	13,596,577	13,596,577	13,596,577
F-Test: Middle School	3.052	1.373	2.023	0.0006
P-Value: Middle School	0.048	0.254	0.133	0.999
F-Test: High School	3.086	2.464	0.528	0.273
P-Value: High School	0.046	0.0856	0.590	0.761

Table C.2.B: Long-term Student Outcomes

	(1) Graduate High School	(2) Enroll College	(3) Enroll: 2-year College	(4) Enroll: 4-year College	(5) Graduate College	(6) Graduate: 2-year College	(7) Graduate: 4-year College	(8) Employed	(9) Income	(10) Income Given Employment
Accept Exposure Effects										
COPS in Schools (CIS)	-0.024* (0.010)	-0.032* (0.013)	-0.045** (0.014)	-0.0009 (0.007)	-0.002 (0.011)	-0.004 (0.008)	0.009 (0.010)	-0.025 (0.022)	-2,553.6*** (650.6)	-2,755.9** (856.0)
Secure Our Schools (SOS)	-0.015 (0.058)	-0.062+ (0.037)	-0.015 (0.036)	-0.0005 (0.019)	-0.038 (0.040)	-0.025 (0.017)	-0.022 (0.033)	0.026 (0.029)	-2,373.7 (1,452.1)	-4,501.5* (2,046.6)
Other Grants	0.003 (0.050)	-0.013 (0.043)	0.006 (0.038)	0.057 (0.071)	0.030 (0.025)	0.005 (0.013)	0.016 (0.021)	0.053 (0.057)	1,772.1 (1,829.5)	768.6 (1,860.7)
Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455
F-Test: Equal	0.159	0.452	1.087	0.347	1.206	1.124	0.524	1.336	2.945	2.464
P-Value: Equal	0.853	0.637	0.338	0.707	0.300	0.325	0.592	0.263	0.0530	0.0856

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Tables 3.4 and ???. Interactions are shown based on the type of COPS grant received, including COPS in Schools (CIS) grants, Secure our Schools (SOS) grants, and other grant partnerships. Comparable interactions are included for application variables (though not shown).

Table C.3: Effects by Grantee Type, School District Police Department vs. Other Grantees

Table C.3.A: Short-term Student Discipline Outcomes

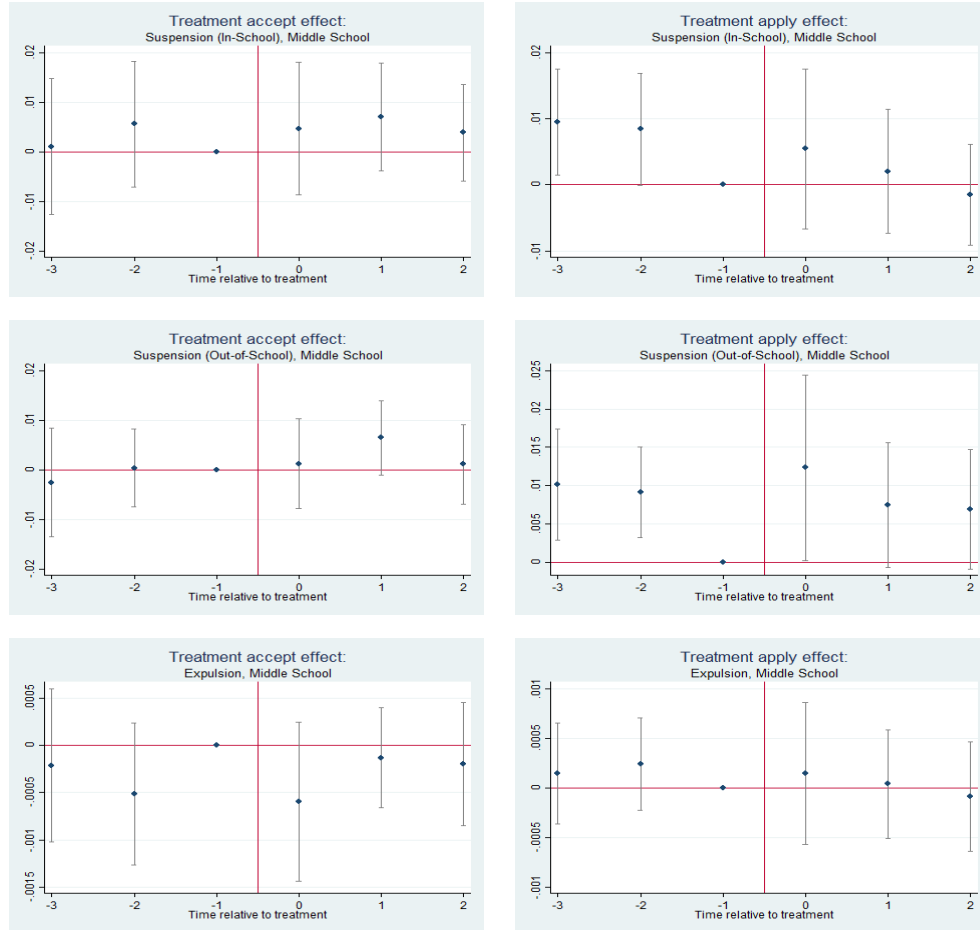
	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
Accept Effects, Middle School				
Not ISD Police	0.013** (0.005)	0.007 (0.005)	0.010* (0.004)	0.0001 (0.0003)
ISD Police	-0.002 (0.010)	-0.010 (0.010)	0.0002 (0.007)	-0.0007 (0.0008)
Accept Effects, High School				
Not ISD Police	-0.005 (0.004)	-0.009+ (0.005)	-0.001 (0.003)	-0.0009*** (0.0002)
ISD Police	-0.012 (0.008)	-0.012 (0.008)	-0.007 (0.005)	-0.0011+ (0.0006)
Observations	13,596,577	13,596,577	13,596,577	13,596,577
F-Test: Middle School	1.810	2.140	1.187	1.048
P-Value: Middle School	0.179	0.144	0.276	0.306
F-Test: High School	0.670	0.114	1.468	0.156
P-Value: High School	0.413	0.736	0.226	0.693

Table C.3.B: Long-term Student Outcomes

	(1) Graduate High School	(2) Enroll College	(3) Enroll: 2-year College	(4) Enroll: 4-year College	(5) Graduate College	(6) Graduate: 2-year College	(7) Graduate: 4-year College	(8) Employed	(9) Income	(10) Income Given Employment
Accept Exposure Effects										
Not ISD Police	-0.047** (0.015)	-0.046** (0.015)	-0.060*** (0.015)	0.002 (0.010)	-0.036** (0.013)	-0.016* (0.007)	-0.027* (0.012)	-0.008 (0.022)	-1,335.6+ (773.1)	-1,407.5+ (820.3)
ISD Police	-0.038 (0.036)	-0.069+ (0.039)	-0.055 (0.044)	-0.027 (0.028)	0.024 (0.037)	0.016 (0.017)	0.019 (0.027)	-0.001 (0.071)	19.2 (2,360.4)	92.9 (2,787.0)
Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455
F-Test: Equal	0.0609	0.287	0.0125	0.773	2.404	2.929	2.335	0.00779	0.367	0.308
P-Value: Equal	0.805	0.592	0.911	0.380	0.121	0.0873	0.127	0.930	0.545	0.579

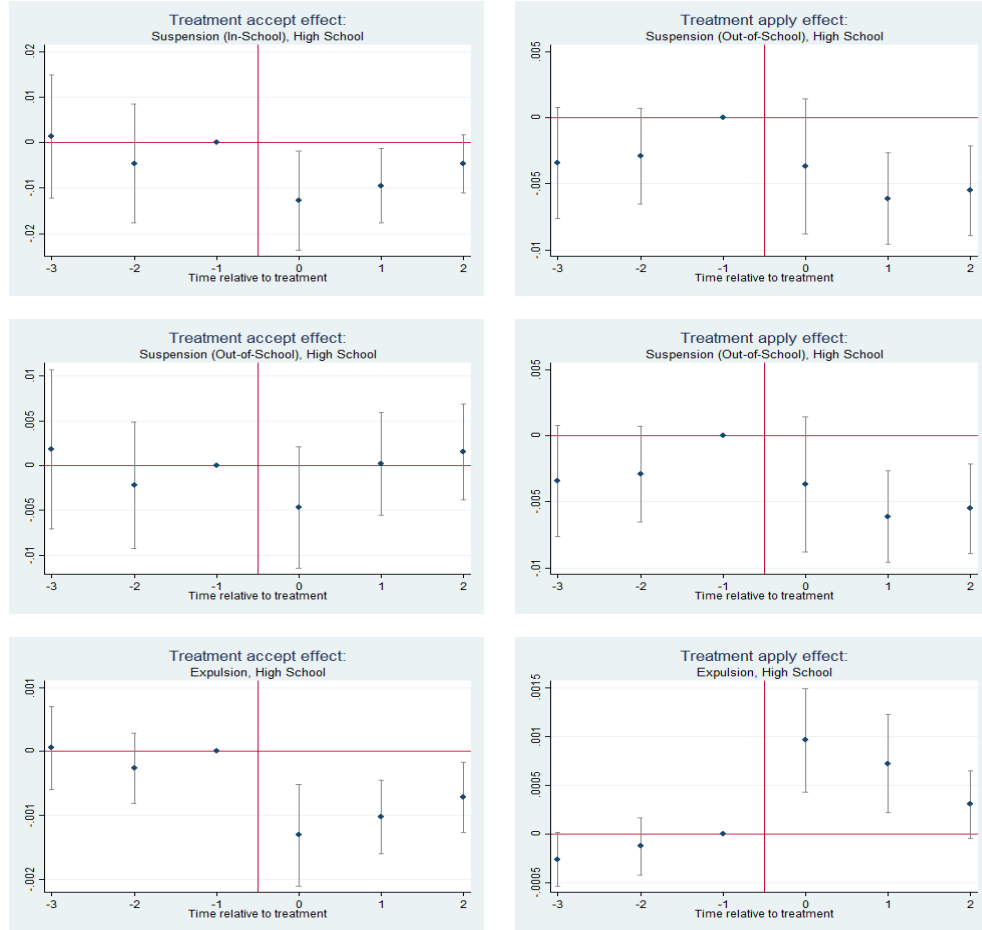
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. Models include a full set of covariates, year fixed effects, grade fixed effects, and district fixed effects as in Tables 3.4 and ???. Interactions are shown based on whether a grantee is an independent school district police department (ISD Police) or is a different type of law enforcement agency (Not ISD Police). Comparable interactions are included for application variables (though not shown).

Figure C.3: Middle School: Timing of Grant Acceptance and Application
Treatment Effects on Student Discipline,



Graphs are produced in a comparable procedure to Figure 3.2. Each of the graphs shows separate coefficient estimates from the same regression. Bars surrounding coefficients represent a 95% confidence interval for each estimate with standard errors clustered at the school district level and the year preceding treatment omitted. Because a school district may receive multiple grants within a sample period, these graphs are created by duplicating the data for each possible treatment year and stacking these data sets to form a "pseudo panel." In each year panel, the designated treatment year is considered over time, and treatments in adjacent years are included as model controls. Due to computational constraints given the size of my data set, these graphs were produced using data collapsed to the school district-grade-year level and weighted by the number of students within these cells. The regressions correspond to the fully specified model, that includes student demographic covariates, student enrollment in a district-grade-year cell, and fixed effects for grade, year, and school district. I include treatments in 2000-2013, so that each centered treatment year has at least one year of observed pre-treatment data.

Figure C.4: High School: Timing of Grant Acceptance and Application Treatment Effects on Student Discipline,



Graphs are produced in a comparable procedure to Figure 3.2. Each of the graphs shows separate coefficient estimates from the same regression. Bars surrounding coefficients represent a 95% confidence interval for each estimate with standard errors clustered at the school district level and the year preceding treatment omitted. Because a school district may receive multiple grants within a sample period, these graphs are created by duplicating the data for each possible treatment year and stacking these data sets to form a "pseudo panel." In each year panel, the designated treatment year is considered over time, and treatments in adjacent years are included as model controls. Due to computational constraints given the size of my data set, these graphs were produced using data collapsed to the school district-grade-year level and weighted by the number of students within these cells. The regressions correspond to the fully specified model, that includes student demographic covariates, student enrollment in a district-grade-year cell, and fixed effects for grade, year, and school district. I include treatments in 2000-2013, so that each centered treatment year has at least one year of observed pre-treatment data.

Table C.4: Table 3.6 with Main Effects: Short-term Student Discipline Outcomes, by Demographic Group

	(1) Disciplinary Action	(2) Suspension (In-School)	(3) Suspension (Out-of-School)	(4) Expulsion
Middle School - Accept Effects				
Economic Disadvantage				
Black	-0.001 (0.009)	-0.006 (0.014)	-0.002 (0.009)	-0.0013 (0.0009)
Hispanic	-0.006 (0.008)	-0.009 (0.008)	-0.004 (0.005)	-0.0005 (0.0006)
White	0.013* (0.006)	0.005 (0.006)	0.016*** (0.005)	-0.0003 (0.0005)
Other Race	0.031** (0.011)	0.026* (0.011)	0.018** (0.006)	0.0003 (0.0006)
No Economic Disadvantage				
Black	0.031*** (0.008)	0.024** (0.009)	0.018** (0.007)	-0.0001 (0.0006)
Hispanic	0.022*** (0.006)	0.014* (0.006)	0.016** (0.005)	0.0006+ (0.0003)
White	0.017*** (0.004)	0.011* (0.005)	0.011*** (0.003)	0.0002 (0.0002)
Other Race	0.024*** (0.007)	0.017* (0.007)	0.014*** (0.004)	0.0003 (0.0004)
High School - Accept Effects				
Economic Disadvantage				
Black	-0.017** (0.005)	-0.021** (0.007)	-0.016** (0.005)	-0.0017** (0.0006)
Hispanic	-0.017** (0.006)	-0.018** (0.007)	-0.003 (0.004)	-0.0008 (0.0005)
White	0.0002 (0.005)	-0.002 (0.006)	0.004 (0.003)	-0.0008+ (0.0005)
Other Race	0.011 (0.010)	0.005 (0.010)	0.012+ (0.006)	-0.0009 (0.0006)
No Economic Disadvantage				
Black	-0.002 (0.009)	-0.007 (0.008)	-0.0001 (0.006)	-0.0008* (0.0004)
Hispanic	0.004 (0.008)	-0.0001 (0.008)	0.003 (0.004)	-0.0012*** (0.0004)
White	0.002 (0.004)	-0.003 (0.005)	0.001 (0.002)	-0.0008*** (0.0002)
Other Race	-0.001 (0.007)	-0.005 (0.006)	0.003 (0.004)	-0.0004 (0.0004)
Main Effects				
Economic Disadvantage				
Black	0.269*** (0.008)	0.220*** (0.008)	0.151*** (0.007)	0.0048*** (0.0004)
Hispanic	0.184*** (0.007)	0.1568*** (0.006)	0.075*** (0.005)	0.0030*** (0.0004)
White	0.168*** (0.005)	0.1426*** (0.004)	0.067*** (0.003)	0.0029*** (0.0003)
Other Race	0.039*** (0.007)	0.0327*** (0.007)	0.014*** (0.004)	0.0009* (0.0004)
No Economic Disadvantage				
Black	0.159*** (0.007)	0.1362*** (0.005)	0.077*** (0.005)	0.0017*** (0.0002)
Hispanic	0.1118*** (0.006)	0.0952*** (0.005)	0.044*** (0.004)	0.0015*** (0.0002)
White	0.046*** (0.004)	0.0369*** (0.004)	0.022*** (0.003)	0.0005* (0.0002)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. This table is equivalent to Table 3.6, but shows additional estimated coefficients for demographic controls in the model, or main effects.

Table C.5: Table 3.7 with Main Effects: Expanded Short-term Student Discipline Outcomes, by Demographic Group

	(1) Conduct Code Violation	(2) Violent Offense	(3) Substance Abuse Offense	(4) Sexual Conduct Offense	(5) Weapon Offense	(6) DAEP	(7) JJAEP	(8) Absent Days	(9) Repeat Grade, Next Year	(10) School Transfer, Next Year	(11) District Transfer, Next Year	(12) Don't Enroll, Next Year
Middle School - Accept Effects												
Economic Disadvantage												
Black	0.006 (0.008)	-0.018* (0.008)	-0.001+ (0.001)	0.0004+ (0.0002)	-0.0008** (0.0003)	0.002 (0.008)	-0.0002 (0.0006)	0.570* (0.226)	-0.002 (0.002)	0.013 (0.008)	-0.013+ (0.006)	-0.004 (0.0035)
Hispanic	-0.005 (0.008)	-0.005 (0.004)	-0.001 (0.001)	0.0001 (0.0001)	-0.0005** (0.0002)	-0.002 (0.004)	-0.0002 (0.0003)	-0.060 (0.161)	-0.0002 (0.003)	0.001 (0.007)	-0.005** (0.002)	0.007*** (0.002)
White	0.014* (0.006)	0.002 (0.002)	0.001 (0.001)	0.0002 (0.0001)	0.00004 (0.0002)	0.009** (0.003)	-0.0003 (0.0004)	0.473*** (0.137)	-0.002 (0.002)	0.003 (0.007)	0.003 (0.003)	0.009*** (0.002)
Other Race	0.034*** (0.009)	0.006* (0.003)	0.002+ (0.001)	0.00001 (0.00004)	0.0004 (0.0002)	0.007* (0.004)	0.0006 (0.0004)	0.243 (0.188)	-0.007* (0.003)	0.004 (0.01)	0.004 (0.004)	-0.001 (0.005)
No Economic Disadvantage												
Black	0.040*** (0.009)	-0.007 (0.005)	0.001* (0.001)	0.0004* (0.0001)	-0.0001 (0.0002)	0.007+ (0.004)	0.0006* (0.0003)	0.505*** (0.126)	-0.005* (0.002)	0.001 (0.008)	0.0025 (0.004)	0.006+ (0.003)
Hispanic	0.027*** (0.006)	0.001 (0.002)	0.001 (0.001)	0.0001* (0.0001)	-0.00001 (0.0001)	0.008*** (0.002)	0.0009** (0.0003)	0.572*** (0.120)	0.008 (0.003)	0.009 (0.009)	0.008* (0.003)	0.009*** (0.002)
White	0.021*** (0.005)	0.007*** (0.001)	0.002*** (0.001)	0.00002 (0.0001)	-0.00009 (0.0002)	0.007** (0.002)	0.0003+ (0.0002)	0.336*** (0.083)	-0.005** (0.002)	0.002 (0.009)	0.005** (0.002)	0.006** (0.002)
Other Race	0.027*** (0.006)	0.011*** (0.002)	0.002*** (0.001)	0.00003 (0.0001)	-0.00001 (0.0002)	0.009** (0.003)	0.0003 (0.0003)	0.308** (0.102)	-0.007* (0.003)	0.014 (0.014)	0.01* (0.005)	0.004 (0.002)
High School - Accept Effects												
Economic Disadvantage												
Black	-0.015** (0.006)	-0.005* (0.002)	-0.003*** (0.001)	-0.0001 (0.0001)	-0.0002 (0.0002)	-0.002 (0.003)	0.0002 (0.0005)	-0.179 (0.259)	0.015*** (0.005)	-0.002 (0.008)	-0.001 (0.002)	-0.011* (0.005)
Hispanic	-0.012+ (0.007)	-0.001 (0.001)	0.001 (0.001)	-0.0001* (0.0004)	-0.0005** (0.0002)	0.002 (0.002)	-0.0003 (0.0003)	-0.485* (0.207)	0.005 (0.005)	-0.008 (0.005)	-0.002 (0.002)	-0.009** (0.003)
White	0.004 (0.005)	0.001 (0.001)	-0.002 (0.001)	-0.0002+ (0.0001)	-0.0003+ (0.0002)	0.006* (0.002)	-0.001*** (0.0003)	-0.115 (0.149)	0.01** (0.003)	-0.007 (0.006)	-0.001 (0.002)	0.0009 (0.002)
Other Race	0.016 (0.010)	0.003* (0.001)	-0.001 (0.001)	-0.0001 (0.0001)	0.0001 (0.0002)	0.0007 (0.003)	-0.0002 (0.0004)	0.262 (0.244)	-0.007 (0.003)	-0.003 (0.011)	0.008* (0.004)	0.007+ (0.004)
No Economic Disadvantage												
Black	0.002 (0.008)	-0.001 (0.002)	-0.002*** (0.001)	0.0001 (0.0002)	-0.0001 (0.0001)	-0.0006 (0.003)	-0.0005 (0.0004)	0.118 (0.222)	0.006+ (0.004)	-0.009 (0.001)	0.001 (0.002)	-0.004 (0.003)
Hispanic	0.009 (0.008)	-0.0003 (0.002)	-0.0004 (0.001)	-0.0001 (0.0001)	-0.0004+ (0.0002)	0.005+ (0.003)	-0.0007** (0.0003)	0.044 (0.191)	0.006 (0.004)	-0.006 (0.007)	0.002 (0.001)	-0.002 (0.003)
White	0.005 (0.004)	0.004*** (0.001)	-0.0003 (0.001)	-0.00002 (0.0004)	-0.0001 (0.0001)	0.003 (0.003)	-0.0004* (0.0002)	-0.094 (0.098)	-0.001 (0.002)	-0.012 (0.008)	0.003** (0.001)	0.002+ (0.001)
Other Race	0.002 (0.007)	0.005*** (0.001)	0.00001 (0.001)	-0.0001 (0.0001)	0.0002 (0.0002)	-0.001 (0.003)	-0.0005 (0.0003)	0.140 (0.156)	-0.007** (0.002)	-0.028* (0.013)	0.002 (0.001)	0.005+ (0.002)
Main Effects												
Economic Disadvantage												
Black	0.245*** (0.008)	0.033*** (0.002)	0.003*** (0.001)	0.001*** (0.0001)	0.0006*** (0.00008)	0.057*** (0.002)	0.0034*** (0.0004)	4.539*** (0.185)	0.046*** (0.002)	0.071*** (0.005)	0.054*** (0.004)	0.026*** (0.002)
Hispanic	0.166*** (0.007)	0.013*** (0.001)	0.008*** (0.001)	0.0003*** (0.0001)	0.0006*** (0.00008)	0.034*** (0.002)	0.0022*** (0.0003)	4.925*** (0.246)	0.044*** (0.003)	0.029*** (0.005)	0.016*** (0.003)	0.021*** (0.002)
White	0.150*** (0.005)	0.009*** (0.001)	0.013*** (0.001)	0.0004*** (0.0001)	0.0007*** (0.00007)	0.032*** (0.001)	0.0018*** (0.0003)	6.184*** (0.134)	0.038*** (0.002)	0.047*** (0.004)	0.057*** (0.002)	0.058*** (0.002)
Other Race	0.032*** (0.007)	0.004 (0.001)	0.001 (0.001)	0.0001+ (0.0001)	0.0002+ (0.0001)	0.007*** (0.002)	0.006+ (0.003)	1.255*** (0.190)	0.008*** (0.002)	0.01*** (0.003)	0.009*** (0.002)	0.013*** (0.003)
No Economic Disadvantage												
Black	0.146*** (0.006)	0.016* (0.001)	0.001* (0.0004)	0.0006*** (0.0001)	0.0003*** (0.00007)	0.024*** (0.002)	0.0013*** (0.0001)	1.214*** (0.158)	0.018*** (0.002)	0.034*** (0.004)	0.026*** (0.003)	0.002 (0.002)
Hispanic	0.101*** (0.006)	0.006*** (0.001)	0.005*** (0.001)	0.0002*** (0.00004)	0.0003*** (0.00007)	0.015*** (0.001)	0.0014*** (0.0003)	3.087*** (0.142)	0.025*** (0.002)	0.019*** (0.004)	0.009*** (0.002)	0.004* (0.002)
White	0.042*** (0.004)	0.002** (0.001)	0.005** (0.0004)	0.0001* (0.00004)	0.0003*** (0.00006)	0.005*** (0.001)	0.0006*** (0.0002)	2.523*** (0.108)	0.009*** (0.001)	0.008** (0.003)	0.006*** (0.001)	0.005** (0.002)
Student-Year Observations	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,596,577	13,292,626	11,684,325	11,684,325	11,684,325	11,684,325

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. This table is equivalent to Table 3.7, but shows additional estimated coefficients for demographic controls in the model, or main effects.

Table C.6: Table 3.8 with Main Effects: Long-term Student Outcomes, by Demographic Group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Graduate High School	Enroll College	Enroll: 2-year College	Enroll: 4-year College	Graduate College	Graduate: 2-year College	Graduate: 4-year College	Employed	Income	Income Given Employment
Acceptance Exposure Effects										
<i>Economic Disadvantage</i>										
Black	-0.001 (0.007)	-0.025*** (0.006)	-0.043*** (0.006)	0.007 (0.006)	0.008 (0.005)	0.005+ (0.003)	0.003 (0.005)	-0.008 (0.012)	228.2 (427.6)	-313.4 (290.1)
Hispanic	0.002 (0.009)	0.003 (0.008)	0.002 (0.009)	0.007 (0.006)	0.012* (0.006)	0.003 (0.003)	0.01+ (0.005)	0.017* (0.009)	714.1*** (198.2)	637.3* (251.0)
White	-0.05*** (0.011)	-0.028*** (0.011)	-0.027*** (0.01)	-0.01 (0.008)	-0.002 (0.006)	0.004 (0.004)	-0.008 (0.006)	-0.027 (0.018)	-14.4 (484.8)	377.4 (819.6)
Other Race	0.009 (0.018)	0.008 (0.032)	0.018 (0.034)	-0.025 (0.016)	-0.003 (0.031)	0.001 (0.011)	-0.01 (0.025)	0.012 (0.017)	-26.6 (405.1)	-266.3 (554.7)
<i>No Economic Disadvantage</i>										
Black	0.003 (0.01)	-0.035* (0.018)	-0.052** (0.018)	0.002 (0.013)	-0.017** (0.006)	-0.002 (0.003)	-0.016* (0.006)	-0.028** (0.009)	-626.1** (222.5)	-322.2 (280.5)
Hispanic	-0.023 (0.014)	-0.03+ (0.017)	-0.031* (0.012)	-0.003 (0.017)	0.005 (0.012)	0.004 (0.005)	0.001 (0.012)	-0.012 (0.008)	-57.5 (280.3)	476.4 (349.1)
White	-0.02 (0.016)	-0.036 (0.023)	-0.019 (0.02)	-0.032+ (0.017)	-0.023 (0.017)	-0.006 (0.004)	-0.021 (0.017)	-0.015 (0.014)	-406.4 (355.3)	123.0 (237.2)
Other Race	-0.001 (0.016)	0.005 (0.023)	0.041* (0.021)	-0.015 (0.029)	-0.007 (0.031)	-0.01 (0.01)	0.002 (0.032)	-0.006 (0.014)	243.1 (458.4)	936.8 (637.1)
Main Effects										
<i>Economic Disadvantage</i>										
Black	-0.146*** (0.009)	-0.162*** (0.011)	-0.149*** (0.012)	-0.172*** (0.01)	-0.221*** (0.01)	-0.039*** (0.004)	-0.193*** (0.009)	0.071*** (0.01)	-2,632.1*** (254.1)	-7,621.7*** (271.3)
Hispanic	-0.113*** (0.01)	-0.243*** (0.012)	-0.169*** (0.013)	-0.266*** (0.01)	-0.227*** (0.01)	-0.018*** (0.004)	-0.217*** (0.009)	0.085*** (0.011)	686.6** (235.2)	-2,373.1*** (324.8)
White	-0.208*** (0.008)	-0.264*** (0.011)	-0.206*** (0.012)	-0.245*** (0.009)	-0.225*** (0.01)	-0.033*** (0.004)	-0.204*** (0.009)	0.032*** (0.01)	-1,469.5*** (237.4)	-4,369.4*** (261.6)
Other Race	-0.035*** (0.008)	-0.024*** (0.01)	0.007*** (0.01)	-0.091*** (0.008)	-0.062*** (0.01)	0.036*** (0.006)	-0.092*** (0.009)	0.032* (0.013)	-415.6 (357.1)	-2,438.4*** (416.8)
<i>No Economic Disadvantage</i>										
Black	-0.02* (0.009)	-0.018+ (0.01)	-0.038** (0.013)	-0.074*** (0.009)	-0.15*** (0.011)	-0.021*** (0.004)	-0.136*** (0.009)	0.116*** (0.009)	-315.3 (235.3)	-5,057.6*** (300.1)
Hispanic	-0.028** (0.009)	-0.075** (0.013)	-0.037** (0.013)	-0.17*** (0.011)	-0.151*** (0.011)	0.004 (0.004)	-0.157*** (0.009)	0.132*** (0.01)	2,248.8*** (242.2)	-1,552.4*** (289.6)
White	-0.009 (0.008)	0.004 (0.01)	0.008 (0.011)	-0.077*** (0.009)	-0.051*** (0.01)	0.011** (0.004)	-0.06*** (0.008)	0.114*** (0.009)	3,270.6*** (226.4)	440.0+ (230.7)
Student Observations	2,514,683	2,514,683	2,514,683	2,514,683	905,506	905,506	905,506	905,506	905,506	550,455

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered at the school district level are shown in parentheses. This table is equivalent to Table ??, but shows additional estimated coefficients for demographic controls in the model, or main effects.

C.2 Data Appendix

Grant Data The grant data for this project was obtained through a FOIA request to the COPS office at the DOJ. The original FOIA request contained information on all accepted and rejected federal COPS grants between 1993-2015. In addition to grants for school police, the total set of grants offered through the COPS office includes hiring grants for traditional police departments, police technology and equipment grants, and other grants for community partnerships or targeted crime initiatives. Both the accepted and rejected grant records included information on the organization applying for a grant, the grant program type, and the project start and end dates. The accepted grant records also included information on the number of officers eligible for funding for hiring grants and the total award size, though this information was often incomplete.

As a first step to cleaning this data, I identified grants for school police in Texas from the total grant files. I did this by extracting grants administered specifically for schools, including COPS in Schools (CIS) hiring grants, Secure our Schools (SOS) security technology grants, and School-based Partnership grants (SBP). Additionally, I extracted grants that independent school district police departments applied for, regardless of their COPS grant program type.

To match the grant start dates to the school calendar, I coded start years to correspond to the current school year if the grant project start date was between September and February, and linked grant projects with start dates between March and August to the following school year (the last third of the school year and

the summer). This strategy assumes that grants received in the last third of the school year and the summer will be utilized in the next year. I next standardized the duration of grants within particular program types, by applying the modal grant project length for a particular grant program for all grants within that program. Grants that last more than one year are coded to correspond to their grant duration for analysis. For example, CIS grants last for 3 years, so if an accepted CIS grant starts in 2000, the variables *Accept_{dt}* and *Apply_{dt}* are coded as 1 for 2000-2002. If a CIS grant was rejected but would have begun in 2000, the variable *Apply_{dt}* is coded as 1 from 2000-2002.

Because school districts do not directly apply for COPS funding, the COPS grant records were manually linked to the appropriate school district or group of school districts. When grants involved independent school district police departments, these links were unique. In cases where the entity applying for a grant was a municipal police department or another organization, web sources and maps were used to match school districts to these grant applying entities. Once a mapping was identified, I collapsed grants to district-years, allowing some school districts to have multiple types of grants or grant rejections and acceptances in the same year (if covered by multiple applying organizations). In all analyses, any acceptance in a given year is coded as an acceptance, even if the school district also was covered by a grant application that was rejected in that year. At the same time, if a grant covers multiple districts in a year, all districts are equivalently coded as acceptances or rejections, depending on the outcome of the grant application. Grant tabulations show in the body of the paper represent "new-school district-grant years," rather

than actual distinct grants distributed to law enforcement agencies.

Student Data The Texas student data used in this project comes from the Texas Education Research Center (ERC), which joins records on primary, secondary schooling from the Texas Education Agency (TEA), post-secondary schooling records from the Texas Higher Education Coordinating Board (THECB), and employment data from unemployment insurance records collected by the Texas Workforce Commission (TWC). For college attendance and graduation outcomes, THECB data covers all public and private post-secondary institutions in Texas but not schools outside of the state. Likewise, employment and earnings records provided by the TWC only cover employment within the state. Student records can be linked across these data sets using unique id numbers created by the ERC.

I created a longitudinal student data set by constructing cohorts of 7th graders that first enrolled in 7th grade in 1999-2006. Throughout the analysis, I assign students to school districts based on their enrollment in the 7th grade. These school district assignments are used to create the district fixed effects as well as to link grant records to students. This design renders the output in the paper as "intent-to-treat" estimates, given that students can change schools and school districts in later years. Separately, I measure student transfers to new schools and new school districts as outcomes to see how student enrollment responds to funding for school police in the original student district that corresponds to 7th grade enrollment. Likewise, failure to enroll in any Texas school district is also included as outcomes.

Time-invariant student demographic characteristics are determined using the modal classification for a particular variable over time, to correct for potential errors. For example, a student may be coded as different races in different years, so the modal classification for each student is used. Multi-race students are not consistently coded across years and districts, so these subgroups are not considered. Instead, multi-racial students are categorized as Black if they are multi-race and Black, then Hispanic, and then other race, prioritizing non-White background in the increasing order of the size of these groups in the sample (other than Asian and other race students who are considered last because they are not distinctly grouped in the analysis). Economic disadvantage, special education, gifted and talented, and LEP status are also assigned the modal value for each student over the panel, though changing values in these characteristics is not common. I exclude student age because grade repeats are included as an outcome and may be impacted by funding for school police, rendering age within a grade level a potentially endogenous regressor. The number of students in a grade-district-year is also included as a control that varies over time within districts.

For short-term discipline outcomes in this study, students are tracked through the 12th grade through 2013, in panels that allow students to repeat up to two grades. Because the analysis allows students to repeat grades, 2011 is the last "on-time" 12th grade year in the sample. High school graduation and college enrollment outcomes are measured within 2 years after an "on time" 12th grade graduation, for all cohorts in the study. College graduation (from both 2 and 4 year schools) are measured within 6 years after an "on time" 12th grade graduation,

or when the typical student is 24 years old. Likewise, employment outcomes are measured during the 6th year after an "on time" high school graduation. Due to limitations in the years of data available, college graduation and employment are measured for the cohorts that enroll in the 7th grade between 1999-2001. The high school graduation variable is constructed to equal 1 if a student graduates from any public Texas high school in the state, regardless of where the student was enrolled in the 7th grade. Similarly, college enrollment and graduation variables are measured as any enrollment or graduation from any post-secondary institution in Texas.

Bibliography

- [1] Daniel Aaronson, Lisa Barrow, and William Sander. Teachers and student achievement in the Chicago Public High Schools. *Journal of Labor Economics*, 25(1):95–135, 2007.
- [2] Lynn Addington. Cops and Cameras: Public School Security as a Policy Response to Columbine. *American Behavioral Scientist*, 52:1426–1446, 2009.
- [3] Anna Aizer and Joseph Doyle. Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges. *The Quarterly Journal of Economics*, 130(2):759–804, 2015.
- [4] Nejat Anbarci and Jungmin Lee. Detecting racial bias in speed discounting: Evidence from speeding tickets in Boston. *International Review of Law and Economics*, 38:11–24, 2014.
- [5] Kate Antonovics and Brian G. Knight. A new look at racial profiling: Evidence from the Boston Police Department. *The Review of Economics and Statistics*, 91(1):163–177, 2009.
- [6] Shamena Anwar and Hanming Fang. An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*, 96(1):127–151, 2006.

- [7] Patrick Bayer, Randi Hjalmarsson, and David Pozen. Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections. *The Quarterly Journal of Economics*, 124(1):105–147, 2009.
- [8] BlueLivesMatter. Opinion: Why do cops always support other officers? <https://bluelivesmatter.blue/>, July 2016.
- [9] Anthony Braga, Andrew Papachristos, and David Hureau. The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly*, 31(4):633–663, 2014.
- [10] Sarah Brayne. Big data surveillance: The case of policing. *American Sociological Review*, (Forthcoming), 2017.
- [11] Ben Brown. Understanding and assessing school police officers: A conceptual and methodological comment. *Journal of Criminal Justice*, 34:591–604, 2006.
- [12] Gregorio Caetano and Vikram Maheshri. Identifying dynamic spillovers of crime with a causal approach to model selection. *Quantitative Economics*, (Forthcoming), 2017.
- [13] A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427, 2008.
- [14] Census. American community survey: Demographic and housing estimates. Technical report, Census Bureau, United States, 2015.

- [15] Aaron Chalfin and Justin McCrary. The effect of police on crime: New evidence from U.S. cities, 1960-2013. *National Bureau of Economic Research*, NBER Working Paper 18815, 2013.
- [16] Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632, 2014.
- [17] Alex Chohlas-Wood, Aliya Merali, Warren Reed, and Theodoros Damoulas. Mining 911 calls in New York City: Temporal patterns, detection and forecasting. *Artificial Intelligence for Cities: Papers from the 2015 AAAI Workshop*, 2015.
- [18] Mark Cohen, Roland Rust, Sara Steen, and Simon Tidd. Willingness-to-pay for crime control programs. *Criminology*, 42(1):89–109, 2004.
- [19] Philip Cook. The clearance rate as a measure of Criminal Justice System effectiveness. *Journal of Public Economics*, 11:135–142, 1979.
- [20] Philip Cook, Max Kapustin, Jens Ludwig, and Douglas Miller. The Effects of COPS Office Funding on Sworn Force Levels, Crime, and Arrest: Evidence from a Regression Discontinuity Design. Technical report, Office of Community Oriented Policing Services, Washington, D.C., 2017.
- [21] COPS. The impact of the economic downturn on American police agencies. *Community Oriented Policing Services, Report of the U.S. Department of Justice (DOJ)*, 2011.

- [22] COPS. COPS in Schools (CIS). *Community Oriented Policing Services Office at the Department of Justice*, 2016. <https://cops.usdoj.gov/default.asp?Item=54>.
- [23] Sergio Correia. A feasible estimator for linear models with multi-way fixed effects. *Working Paper*, 2016.
- [24] Chris Curran. Estimating the Effect of State Zero Tolerance Laws on Exclusionary Discipline, Racial Discipline Gaps, and Student Behavior. *Educational Evaluation and Policy Analysis*, 20(10):1–22, 2016.
- [25] Gregory DeAngelo and Benjamin Hansen. Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, 6(2):231–257, 2014.
- [26] Department of Justice. Supporting Federal, State, Local and Tribal Law Enforcement. Memorandum, Office of the Attorney General, Department of Justice, United States, March 2017.
- [27] Matthew Desmond, Andrew Papachristos, and David Kirk. Police violence and citizen crime reporting in the Black community. *American Sociological Review*, 81(5):857–876, 2016.
- [28] Jill DeVoe, Katharin Peter, Phillip Kaufman, Sally Ruddy, Amanda Miller, Mike Planty, Thomas Snyder, and Michael Rand. Indicators of School Crime and Safety: 2003. *National Center for Education Statistics at U.S. Department*

of Education and Bureau of Justice Statistics at U.S. Department of Justice Report, 2003.

- [29] Rafael Di Tella and Ernesto Schargrodsky. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133, 2004.
- [30] John Donohue. Assessing the relative benefits of incarceration: The overall change over the previous decades and the benefits on the margin. In Stephen Raphael and Michael Stoll, editors, *Do Prisons Make Us Safer? The Benefits and Costs of the Prison Boom*. Russell Sage Foundation, New York, NY, 2009.
- [31] Mirko Draca, Stephen Machin, and Robert Witt. Panic on the streets of London: Police, crime, and the July 2005 terror attacks. *American Economic Review*, 101(5):2157–2181, 2011.
- [32] Jennifer Eberhardt, Phillip Goff, Valerie Purdie, and Paul Davies. Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, 87(6):876–893, 2004.
- [33] EOP. Report: The Continuing Need to Rethink Discipline. *Executive Office of the President Report*, 2016.
- [34] William N. Evans and Emily G. Owens. COPS and crime. *Journal of Public Economics*, 91(1-2):181–201, 2007.
- [35] Tony Fabelo, Michael Thompson, Martha Plotkin, Dottie Carmichael, Miner Marchbanks, and Eric Booth. Breaking Schools’ Rules: A Statewide Study

of How School Discipline Relates to Students' Success and Juvenile Justice Involvement. *Justice Center at The Council of State Governments and Public Policy Research Institute*, 2011.

- [36] FBI. Crime in the United States: Offenses Known to Law Enforcement. Technical report, Federal Bureau of Investigation, Department of Justice, United States, 2015.
- [37] Benjamin Fisher and Emily Hennessy. School Resource Officers and Exclusionary Discipline in U.S. High Schools: A Systemic Review and Meta-analysis. *Adolescent Research Review*, 1:217–233, 2016.
- [38] Deborah Fowler, Rebecca Lightsey, Janis Monger, and Elyshia Aseltine. Texas' School-to-Prison Pipeline: Ticketing, Arrest and Use of Force in Schools, How the Myth of the "Blackboard Jungle" Reshaped School Disciplinary Policy. *Texas Appleseed Report*, 2010.
- [39] Kathleen Frydl and Wesley Skogan. *Fairness and Effectiveness in Policing: The Evidence*. National Academies Press, 2004.
- [40] Roland Fryer. An empirical analysis of racial differences in police use of force. *National Bureau of Economic Research*, 22399, 2016.
- [41] GAO. Technical assessment of Zhao and Thurman's 2001 evaluation of the effects of COPS grants on crime. *Government Accountability Office*, GAO-03-867R, 2003.

- [42] GAO. Community policing grants: COPS grants were a modest contributor to declines in crime in the 1990s. *Government Accountability Office*, GAO-06-104, 2005.
- [43] Luis Garicano and Paul Heaton. Information technology, organization and productivity in the public sector: Evidence from police departments. *Journal of Labor Economics*, 29(1):167–201, 2010.
- [44] Simen Gaure. lfe: Linear group fixed effects. *The R Journal*, 5(2):104–116, 2013.
- [45] Simen Gaure. OLS with multiple high dimensional category variables. *Computational Statistics and Data Analysis*, 66(8-18), 2013.
- [46] Andrew Gelman, Jeffrey Fagan, and Alex Kiss. An analysis of the New York City police department’s "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, 102(479):813–823, 2007.
- [47] Filipe Goncalves and Steven Mello. Racial bias in policing: Evidence from speeding tickets. *Working Paper*, 2017.
- [48] Jeffrey Grogger and Greg Ridgeway. Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association*, 101(475):878–887, 2006.
- [49] Cassandra Guarino, Michelle Maxfield, Mark Reckase, Paul Thompson, and Jeffrey Wooldridge. An evaluation of Empirical Bayes’ estimation of value-

- added teacher performance measures. *Journal of Educational and Behavioral Statistics*, 40(2):190–222, 2015.
- [50] Paulo Guimaraes and Pedro Portugal. A simple feasible procedure to fit models with high-dimensional fixed effects. *The Stata Journal*, 10(4):628–649, 2010.
- [51] Stefan Harrendorf, Markku Heiskanen, and Stephen Malby. International Statistics on Crime and Justice. *United Nations Office on Drugs and Crime*, 2010.
- [52] Christopher Haugh. How the Dallas Police Department Reformed Itself. *The Atlantic*, July 9 2016.
- [53] Angela Hawken and Mark Kleiman. Managing drug involved probationers with swift and certain sanctions: Evaluating Hawaii’s HOPE. *Department of Justice*, 2009.
- [54] James Hines and Richard Thaler. Anomalies: The Flypaper Effect. *Journal of Economic Perspectives*, 9(4):217–226, 1995.
- [55] Randi Hjalmarsson. Criminal justice involvement and high school completion. *Journal of Urban Economics*, 63:613–630, 2008.
- [56] William C. Horrace and Shawn M. Rohlin. How dark is dark? bright lights, big city, racial profiling. *The Review of Economics and Statistics*, 98(2):226–232, 2016.

- [57] Nathan James and Gail McCallion. School Resource Officers: Law Enforcement Officers in Public Schools. *Congressional Research Service Report to Congress*, 2013.
- [58] Elizabeth Joh. Surveillance discretion: Automated suspicion, big data, and policing. *Harvard Law and Policy Review*, 10:15–42, 2016.
- [59] Rucker Johnson. Ever-increasing levels of parental incarceration and the consequences for children. In Stephen Raphael and Michael Stoll, editors, *Do Prisons Make Us Safer?: The Benefits and Costs of the Prison Boom*, pages 177–206. Russell Sage Foundation, New York, NY, 2009.
- [60] Rucker Johnson and Stephen Raphael. How Much Crime Reduction Does the Marginal Prisoner Buy? *Journal of Law & Economics*, 55(2), 2012.
- [61] Jeffrey Jones. In U.S., confidence in police lowest in 22 years. Technical report, Gallup, June 2015.
- [62] Thomas Kane and Douglas Staiger. Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research*, 14607, 2008.
- [63] David Kirk and Robert Sampson. Crime and the Production of Safe Schools. In Greg Duncan and Richard Murnane, editors, *Whither Opportunity? Rising Inequality, Schools, and Children’s Life Chances*. Russell Sage Foundation, New York, NY, 2011.

- [64] Jonathan Klick and Alexander Tabarrok. Using terror alert levels to estimate the effect of police on crime. *Journal of Law and Economics*, 48(1):267–279, 2005.
- [65] John Knowles, Nicola Persico, and Petra Todd. Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, 109(1):203–229, 2001.
- [66] Cory Koedel, Kata Mihaly, and Jonah E. Rockoff. Value-added modeling: A review. *Economics of Education Review*, 47:180–195, 2015.
- [67] Nancy La Vigne, Jocelyn Fontaine, and Anamika Dwivedi. How do people in high-crime, low-income communities view the police? Technical report, Urban Institute, February 2017.
- [68] David Lee and Justin McCrary. The deterrence effect of prison: Dynamic theory and evidence. *Princeton University, Department of Economics*, Working Paper 1168, 2009.
- [69] Jae-Seung Lee, Jonathan Lee, and Larry Hoover. What conditions affect police response time? Examining situational and neighborhood factors. *Police Quarterly*, 20(1):61–80, 2017.
- [70] Steven Levitt. Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3):270–290, 1997.
- [71] Steven Levitt. Using electoral cycles in police hiring to estimate the effect of police on crime: Reply. *American Economic Review*, 92(4):1244–1250, 2002.

- [72] Ming-Jen Lin. More police, less crime: Evidence from U.S. state data. *International Review of Law and Economics*, 29:73–80, 2009.
- [73] Lance Lochner. Individual perceptions of the Criminal Justice System. *American Economic Review*, 97(1):444–460, 2007.
- [74] Alexandre Mas and Enrico Moretti. Peers at work. *The American Economic Review*, 99(1):112–145, 2009.
- [75] Stephen Mastrofski. Controlling street-level police discretion. *The Annals of the American Academy of Political and Social Science*, 593:100–118, 2004.
- [76] Justin. McCrary. Using electoral cycles in police hiring to estimate the effect of police on crime: Comment. *American Economic Review*, 92(4):1236–1243, 2002.
- [77] Steven Mello. Police and Crime: Evidence from COPS 2.0. *Working Paper*, 2016.
- [78] Ted Miller, Mark Cohen, and Brian Wiersema. Victim costs and consequences: A new look. *National Institute of Justice Research Report*, 1996.
- [79] Carl Morris. Parametric Empirical Bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78(381):47–55, 1983.
- [80] Michael Mueller-Smith. The criminal and labor market impacts of incarceration. *Working Paper*, 2015.

- [81] Richard Murnane. U.S. High School Graduation Rates: Patterns and Explanations. *Journal of Economic Literature*, 51(2):370–422, 2013.
- [82] Chongmin Na and Denise Gottfredson. Police Officers in Schools: Effects on School Crime and the Processing of Offending Behaviors. *Justice Quarterly*, pages 1–32, 2011.
- [83] Daniel S. Nagin, Robert M. Solow, and Cynthia Lum. Deterrence, criminal opportunities and police. *Criminology*, 53(1):74–100, 2015.
- [84] NEA. Rankings and Estimates: Rankings of the States 2015 and Estimates of School Statistics 2016. *National Education Association Research Report*, 2016.
- [85] Ernest Nickels. A note on the status of discretion in police research. *Journal of Criminal Justice*, 35:570–578, 2007.
- [86] Kathleen Nolan. *Police in the Hallways: Discipline in an Urban High School*. University of Minnesota Press, Minneapolis, MN, 2011.
- [87] Wallace Oates. Toward a Second-Generation Theory of Fiscal Federalism. *International Tax and Public Finance*, 12:349–373, 2005.
- [88] OIG. Police hiring and redeployment grants: Summary of audit findings and recommendations, October 1996-September 1998. *Office of the Inspector General Special Report, Department of Justice*, 99-14, 1999.
- [89] Emily Owens. COPS and cuffs. In Olivier Marie Philip J. Cook, Stephen Machin and Giovanni Mastrobuoni, editors, *Lessons from the Eco-*

nomics of Crime: What Reduces Offending? MIT Press, Cambridge, MA, 2013.

- [90] Devah Pager, Bruce Western, and Naomi Sugie. Sequencing disadvantage: Barriers to employment facing young Black and White men with criminal records. *The Annals of the Academy of Political and Social Science*, 623(1):195–213, 2009.
- [91] PERF. Local police rushing to apply for stimulus bill grants. *Police Executive Research Forum Newsletter*, 23(3), 2009.
- [92] The. Pew Charitable Trusts. Collateral costs: Incarceration’s effect on economic mobility. Technical report, The Pew Charitable Trusts., 2010.
- [93] Bill Schneider. Do Americans trust their cops to be fair and just? New poll contains surprises. *Reuters*, January 2015.
- [94] David Silver. Haste or waste? Peer pressure and the distribution of marginal returns to health care. *Working Paper*, 2016.
- [95] Texas. Senate Bill No. 393. *Texas Legislature*, 2013.
- [96] Texas. House Bill No. 2684. *Texas Legislature*, 2015.
- [97] Tasha Tsiaperas. Top Dallas cops take classes to learn how to fight bias. *The Dallas Morning News*, June 2016.
- [98] Tom Tyler, Phillip Goff, and Robert MacCoun. The impact of psychological science on policing in the united states: Procedural justice, legitimacy,

- and effective law enforcement. *Psychological Science in the Public Interest*, 16(3):75–109, 2015.
- [99] Roy Walmsley. World Prison Population List (11th Edition). *Institute for Criminal Policy Research*, 2016.
- [100] David Weisburd, Elizabeth Groff, Greg Jones, Breanne Cave, Karen Amendola, Sue-Ming Yang, and Rupert Emison. The Dallas patrol management experiment: Can AVL technologies be used to harness unallocated patrol time for crime prevention? *Journal of Experimental Criminology*, 11(367-391), 2015.
- [101] David Weisburd and Cody Telep. Hot spots policing: What we know and what we need to know. *Journal of Contemporary Criminal Justice*, 30(2):200–220, 2014.
- [102] Sarit Weisburd. Police presence, rapid response rates, and crime prevention. *Working Paper*, 2016.
- [103] Emily Weisburst. Safety in Police Numbers: Evidence of Police Effectiveness from Federal COPS Grant Applications. *Working Paper*, 2017. Available at SSRN: <https://ssrn.com/abstract=2845099>.
- [104] Jeremy West. Racial bias in police investigations. *Working Paper*, 2015.
- [105] Anlan Zhang, Lauren Musu-Gillette, and Barbara Oudekerk. Indicators of School Crime and Safety: 2015. *National Center for Education Statistics at U.S. Department of Education and Bureau of Justice Statistics at U.S. Department of Justice Report*, 2016.

- [106] Jihong Zhao, Matthew Scheider, and Quint Thurman. Funding community policing to reduce crime: Have COPS grants made a difference? *Criminology and Public Policy*, 2(1):7–32, 2002.

Vita

Emily Karen Weisburst received her Bachelor of Arts from Dartmouth College in 2008. Prior to graduate school she worked as an associate at the management consulting firm L.E.K. Consulting and as a research assistant for Professor Paul Gompers at Harvard Business School. During graduate school, Emily worked as a Staff Economist at the Council of Economic Advisers in the Executive Office of the President and as a research associate for the RAND Corporation on joint projects with the Texas Higher Education Coordinating Board. Emily also received the NAED Spencer Dissertation Fellowship to support her research on the impact of increasing police presence in public schools on student disciplinary outcomes and educational attainment in Texas. Emily's research focuses on topics in labor economics and public finance, including criminal justice and education.

Permanent address: 105 W. 51st St., Apt. 5204
Austin, Texas 78751

This dissertation was typeset with L^AT_EX[†] by the author.

[†]L^AT_EX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's T_EX Program.